DOCUMENT RESUME

ED 349 828 FL 020 647

AUTHOR Secada, Walter

TITLE Evaluating the Mathematics Education of Limited English Proficient Students in a Time of Educational Change.

PUB DATE Aug 92

NOTE 64p.; In: Focus on Evaluation and Measurement. Volumes 1 and 2. Proceedings of the National Research Symposium on Limited English Proficient Student Issues (2nd, Washington, DC, September 4-6, 1991); see FL 020 630.

PUB TYPE Viewpoints (Opinion/Position Papers, Essays, etc.) (120)

EDRS PRICE MF01/PC03 Plus Postage.

DESCRIPTORS Curriculum Design; *Educational Change; *Educational Objectives; Elementary Secondary Education; *Limited English Speaking; *Mathematics Education; *Mathematics Instruction; *Program Evaluation

ABSTRACT

Program evaluation and related research has come a very long way from the quasi-experiment to where it is now seen as having many functions, as being grounded in a range of theoretical positions, and as drawing from a variety of possible methodologies. This paper focuses on evaluating the mathematics programs of limited-English-proficient (LEP) students in a time of educational change. Specific sections address the following: bilingual education program goals; the measurement of goals; the "taken-for-granted" status in current evaluation practice; the mathematics curriculum; mathematics learning and thinking; and mathematics instruction. (VWL)

Evaluating the Mathematics Education of Limited English Proficient Students in a Time of Educational Change

Walter Secada
University of Wisconsin, Madison

Bilingual-Education-Program Evaluation: Current Practice

Program evaluation and related research have come a very long way from the quasi-experiment as formalized by Campbell and Stanley (1966; Cook & Campbell, 1979) to where program evaluation is now seen as having many functions, as being grounded in a range of theoretical positions, and as drawing from a variety of possible methodologies (Cook & Shadish, 1986; Cronbach, 1980; Lindblom & Cohen, 1979). In practice, however, the evaluation of bilingual education programs has not strayed very far from its original, basic question: Does the program work better than not having the program? Or, Does the program work better than having a particular, alternative program?[1] At one time, the law that provided federal funds for bilingual education required districts to compare performance by students who were in the program to performance by students who were not. This has been the minimal question that the evaluation of federally funded programs should try to answer.

Regardless of this narrow focus in bilingual-education-program evaluation, it has been de rigueur to bemoan the quality of evaluations that have been produced by federally funded projects. On this point, sympathizers, critics, and people who are neutral about bilingual education <u>all</u> seem to agree (Baker & De Kanter, 1983; Boruch & Cordray, 1980; Willig, 1985).

Elsewhere, I have speculated on some of the reasons for these two problems with current practice in bilingual-education-program evaluation: (a) the failure to move beyond a very narrow set of questions to other questions that are no less interesting and that are, in many ways, more important to local stakeholders; and (b) the failure to meet technical standards of rigor. This is not to claim that there have been no advances in the field. New models for program evaluation, the best known being the gap-reduction model (Tallmadge, Lam, & Gamel, 1987a, 1987b) have been developed. And, federally funded large-scale evaluations of bilingual education have come a very long way from the AIR Report (Danoff, 1978; Danoff, Coles, McLaughlin, & Reynolds, 1977-1978) when there were no efforts to ensure prior-to-treatment comparability of the comparison groups or

to document the fidelity of programs to their descriptions. Though many people on all sides of the "effectiveness debate" may not be satisfied with the conclusions of the Longitudinal Study (Ramirez, Pasta, Yuen, Billings, & Ramey, 1991; Ramirez, Yuen, & Ramey, 1991; Ramirez, Yuen, Ramey & Pasta, 1991), it did document fidelity of treatment and it ensured comparability of groups. That study has also moved the field of program evaluation forward in many other ways: for example, it served as a testing ground for new statistical methods like hierarchical linear modeling (HLM).

Regardless of these developments, there have been at least two constant foci of debate in bilingual-education-program evaluation on federal and local scales. These debates have been around the goals of the program and the sorts of evidence that evaluation can provide.

Bilingual Education Program Goals

Over the years, there has been quite a bit of debate about the range of goals that are appropriate for programs of bilingual education. At the start of the federal funding initiatives, from the late 1960s and into the early 1970s, this debate was couched in terms of two poles that used a variety of terms: assimilation and monoculturalism versus pluralism and biculturalism, the development of English and of English literacy versus native language maintenance, the development of balanced bilingualism, and biliteracy (Andersson & Boyer, 1978; Mackey & Beebe, 1977; Stein, 1986). Pluralist views on the purposes of the program came under concerted attack almost as soon as they were articulated (e.g., Epstein, 1977), and the AIR Report (Danoff et al., 1977-1978) found that such programs did not enhance elementary-school Hispanic children's achievement (in English) better than if there had been no program in place.[2] The mid 1970s was a time of retreat from the purported excesses of the late 1960s, among them cultural pluralism. Thus, the federal-funding program has come to be sharply defined around two goals: the development of English language skills by LEP students and the development of their academic skills so as not to fall progressively behind their English-proficient peers (Secada, 1990a; Stein, 1986).

In recent debates about the goals for bilingual education, some authors have written as if the federal government were mandating a single approach or as if the only goal of the program were to develop English-language skills (see Baker & de Kanter, 1983; Government Accounting Office, 1987). Eleanor Chelimsky, director of the Program Evaluation and Methodology Division of the GAO, argued against this overly narrow specification of goals when she testified before Congress:

> To say, first of all, that there is a method mandated in the act when, in fact, the act says we will use native language to the degree that is necessary -- that is all it says....The same is true for the business of the two goals. To say the act only has one goal, teaching English, is to ignore the other goal of the act which has to do with keeping people up to date in all their subjects. (Reauthorization..., 1987, p. 30)

Other writers have argued that, granting the transitional and assimilationist bent of the above goals, they are too modest. For example, in reviewing the research literature on culturally-diverse populations and mathematics achievement, I noted that there should be at least three goals for intervention programs such as Chapter 1 and bilingual education: (a) improved achievement (beyond what would have occurred without the program); (b) a closing of the achievement gap between the population of interest and the so-called mainstream; and (c) long-term effects wherein the gap, once it is closed, remains closed (Secada, in press). The first of these goals is clearly a goal for bilingual education. The second appears, at least tacitly, in status studies that report the mathematics achievement of diverse learning populations compared to one another (reviewed in Secada, in press), in the gap-reduction model (Tallmadge et al., 1987a, 1987b), and in research designs like that of the Longitudinal Study (Ramirez, Pasta, Yuen, Billings, & Ramey, 1991). Though the third goal has been an explicit part of longitudinal studies like the Sustaining Effects Study and other evaluations of Chapter 1 (Kennedy, Jung, & Orland, 1986). I have been unable to find any evidence that the third goal -- long-lasting closure of the achievement gap -- has been considered in the design or the evaluation of bilingual education programs.

Orum (1983) argued for long-term and for nonacademic goals in bilingual education: reduced dropout rates and an increase in successful school-completion, transition from high school to postsecondary education or to the workplace, and staying at grade level. Christina Bratt-Paulston (1980) also has argued that the goals for bilingual education should be long range and that they should include out-of-school outcomes. Among her recommended indicators of success are:

> employment figures upon leaving school, figures on drug addiction and alcoholism, suicide rates, and personality disorders, that is, indicators which measure the social pathology which accompanies social injustice rather than attempts at efficient language teaching -- although programs are that too (p. 41).

There is wisdom in these recommendations, not only because of the vision of schooling that they propose but also because the payoffs for programs such as bilingual education may, in fact, be long term.

The case of Head Start is illustrative in this regard. Initially, Head Start's goals were short-term and cognitive. And on those grounds, that program fell into deep trouble, much as has been the case for bilingual education. It was on the basis of the long-term, and especially the out-of-school, outcomes of Head Start that it finally achieved the widespread social science and political support that it currently has (Stallings & Stipek, 1986; White & Buka, 1987).

Measurement of Goals

The measurement of bilingual-education-program goals, especially of its academic goals, usually has been translated to mean academic achievement. Typically, as in the case of Chapter 1 evaluations, reading and mathematics are the subjects for which academic achievement information has been gathered.

There have been some debates about the language of the achievement tests that are administered. Some writers have argued that, since the ultimate goal is for students to function in an all-English-speaking setting, achievement should be measured only via English language tests (Baker & de Kanter, 1983; Danoff et al., 1987-1978). Others have argued that, even though the eventual goal is to function in an all-English setting, achievement in either language should be measured in order to get as complete a picture as we can of students' actual learning of content (Willig, 1985; Ramirez, Pasta, et al., 1991). As a proxy for achievement, large-scale studies involving bilingual populations also have used indicators of engaged time on task (e.g., Tikunoff, 1985).

In mathematics achievement, the AIR Study (Danoff et al., 1977-1978) found that only in fourth-grade mathematics achievement did children enrolled in bilingual programs outperform children who were in neighboring school districts and were not enrolled in such programs. In their narrative review of the bilingual-education-program evaluation research, Baker and De Kanter (1983) found uneven effects of programs on mathematics achievement. However, in her meta-analysis of a subset of the Baker and de Kanter studies, Willig (1985) found that children enrolled in bilingual programs outperformed control children on standardized tests of mathematics achievement, whether those tests were administered in English or in Spanish. Interestingly, Willig also found that the better the technical quality of a study -- e.g., if it used random assignment of students -- the more likely it was that the evaluation would show favorable results for the program.

Ramirez and his colleagues did not conduct a direct comparison of various program models[3] against each other due to confounding school-level with program-level effects. In an effort to circumvent those problems, Ramirez et al. compared how well students in each

program performed against the norming populations for the standardized tests that were administered. Between kindergarten and first grade, children in all three programs grew more quickly on an English-language standardized test of mathematics than did the norming populations (see Ramirez, Yuen, & Ramey, 1991, Figures 7, 8, & 9). Between first and third grades, children in all three programs kept pace with the norming population (Figures 10, 11, & 12).

Next, Ramirez and his associates compared the growth in mathematics achievement among students who were enrolled in late-exit bilingual education programs, and had experienced different amounts of their native language (Spanish) over the course of their elementary school years. Students who experienced the most substantial and the most consistent use of Spanish began below national norms but grew the most in mathematics achievement.

> Students in site E, who were provided with substantial instruction in their primary language and a slow phasing in of English instruction over time, consistently realized the greatest growth in mathematics skills, faster than [the] norming population. Students in site D, who were exposed to a consistent proportion of instruction in their primary language (approximately 40 percent), realized growth in mathematics that was equal to [the] norming population. Noteworthy is that after covariates were considered, there was no difference in achievement of students in sites D and E, <u>although students in site E had more stress in their environment and fewer resources than site D students</u>. (Ramirez, Yuen, & Ramey, 1991, p. 33, emphasis added)

In other words, even though the students in one site lived in greater poverty and experienced more of what Ramirez (in personal communication) has termed the stresses of urban life (e.g., crime), they exceeded the norming population's growth and kept pace with a relatively more advantaged population. The tenor of Ramirez et al.'s observations leave little doubt that they ascribe this to the students' receiving substantial amounts of instruction via their native language, Spanish. Consider their observations about the third school in this sample:

> It appears that students in site G who received about 40 percent of their instruction in their primary language in kindergarten and first grade, but were then abruptly moved into almost exclusive instruction in English (comparable to that provided to early-exit and immersion strategy programs), experienced a <u>marked decrease</u> in growth in mathematics skills over time relative to [the] norming population. It seems that these students lost ground...paralleling what is commonly observed for disadvantaged students in the general population. (p. 33, emphasis added)

Surprisingly, though Ramirez and his associates collected achievement data via the students' native language of Spanish, they failed to report aggregate achievement data in Spanish and they did not analyze those data as they analyzed their English-language achievement data.

Thus, the best evidence that we have at this moment suggests that the use of children's native language -- at least for Spanish-speaking children -- for instruction in mathematics is more efficacious than instruction all in English. Moreover, the Ramirez study suggests that the more substantial and consistent the use of a child's native language during the primary-school years, the greater that child's growth will be -- up to the point where the gap between LEP and an English-proficient norming population actually decreases.

Omitted from most bilingual-education-program evaluations are other indicators of academic growth and whether or not, on those indicators, LEP students function similarly to their English-proficient peers. Continued course taking in mathematics should be one such concern (Chipman & Thomas, 1987; Oakes, 1990a, 1990b). Though achievement is important, the continued taking of mathematics courses is at least equally important since, regardless of achievement, one cannot take advanced courses without having taken earlier courses. In their study of the determinants of mathematics course taking by various ethnolinguistic populations in the High School and Beyond (HSB) data base,[4] Myers and Milne (1982) found differential patterns of course offerings and of course taking by high school males and females. We need to understand the reasons for such patterns and what we can do to encourage LEP students to take more mathematics courses in high school.

One reason that most program evaluations fail to attend to non-achievement indicators may be that most bilingual education programs are in elementary school where everyone takes the same mathematics course -- arithmetic -- and course taking does not seem to be an issue. But by junior high school, course taking is becoming optional and, beginning at these grades, it should be (but has not been) a concern.

The Taken-for-Granted in Current Evaluation Practice

Compensatory education was established with the idea of providing students the experiences and the skills that purportedly had been denied to them because of their culturally or linguistically impoverished upbringing (Kantor, 1991; Stein, 1986). Consistent with this belief, evaluation did not question the nature or the quality of curriculum or instruction that these students received. Curriculum and instruction were assumed as given.

Over the years, many writers have rejected such notions of deprivation that undergird the Great Society's Compensatory Education thrust (Kantor, 1991). But the programs that grew from that thrust and many of the assumptions that undergird those programs (and their evaluations) persist.

The Mathematics Curriculum

This general acceptance of the school mathematics curriculum is reflected in current bilingual-education-program research and evaluation practice. I have never seen efforts to document whether or not curricular objectives or materials are different for LEP versus mainstream students. In my own informal observations, however, I have noticed that, when a program for LEP students assumes the responsibility for the mathematics instruction of LEP students, the curriculum is very much focused on computational skills, and instruction tends to be individualized seatwork on pages and pages of worksheets. Mathematics instruction for Chapter 1 students (Kennedy, Jung, & Orland, 1986) or for students enrolled in low track courses (Oakes, 1990a, 1990b) can be similarly characterized.

Efforts to adapt the mathematics curriculum that LEP students receive have come about, mainly through content-based, English-as-a-second-language (ESL) approaches. The goal of these efforts is to develop English language skills through student engagement in mathematics, science, and social studies (Cantoni-Harvey, 1987; Crandall, 1987; Mohan, 1986). These approaches include a structural-linguistic analysis of what has been termed the mathematics register, and they tie that analysis to recommended goals for combining the teaching of mathematics with the teaching of English (Crandall, Dale, Rhodes, & Spanos, 1987; Dale & Cuevas, 1987; Spanos, Rhodes, Dale & Crandall, 1988).

O'Malley and Chamot (1990) have conducted an extensive series of studies documenting the learning strategies used by second-language learners as they learned their second languages (English being among the languages of interest), and for in-school populations, as they learned academic subjects such as mathematics. The results of their studies have included curriculum materials (Chamot & O'Malley, 1988) that try to combine second-language-learning and mathematics-learning.

For both of these approaches, content-based-ESL and language-learning-strategies, mathematics remains constant. There are no questions about its goals and objectives, nor about the adequacy of extant curriculum to meet those goals.

Students' Mathematics Learning and Thinking

Both content-based-ESL and learning-strategies approaches for teaching LEP students might help provide insights into how bilingual students learn mathematics. They entail at least tacit critiques that current mathematics teaching fails to match how people learn a second language and that it may not match -- how LEP students actually learn mathematics. For example, one might use the structural linguistic analyses provided by Crandall and her colleagues (Crandall, Dale, Rhodes, & Spanos, in press; Dale & Cuevas, 1987; Spanos, Rhodes, Dale, & Crandall, 1988) to argue that the reason LEP students do not achieve as well in mathematics as their English-proficient peers is that they lack knowledge of the mathematics register (Orr, 1987, makes a similar claim for students who speak Black English Vernacular). Unfortunately, there is no evidence that English-proficient students have any better grasp of that same register. Were such evidence forthcoming, it would provide a linguistic basis for looking at the school mathematics curriculum.

Carpenter (1985) has argued that, as early as first grade, the school mathematics curriculum ignores the rich stores of informal mathematical (as opposed to linguistic) knowledge that children bring to school. That mismatch, according to Carpenter, lays the foundation for widespread failure and disenchantment with mathematics among older children. Unlike other claims about children, Carpenter's is an argument based on competence -- children enter school competent in mathematical reasoning; the schools ignore that competence; and hence, the typical result of schooling is incompetence in mathematics. A similar case might be built vis-à-vis bilingual students.

There is a common folklore that bilingual students cannot solve arithmetic word problems and that the best we can hope for is to provide them with key words and other tricks for solving such problems. But in my work, I have found that first grade Hispanic bilingual children can solve many of the same word problems that have been used in studies involving monolingual children (Secada, 1991a). Moreover, I have found that competence in solving arithmetic word problems varies as a function of children's proficiency in the language in which they are assessed and also in degree of bilingualism when that language proficiency is assessed qua mathematical language.

Finding Out/Descubrimiento (FO/D; De Avila, Cohen, & Intili, 1982; De Avila, Duncan, & Navarrete, 1987) seems to have been developed along lines that combine what was known about concept formation and second language learning. Like Carpenter's argument, it is based on the tacit assumption that LEP students have more capacity than they are usually credited with. But FO/D extends

Carpenter's argument to include both academic and linguistic competence. Cheche Konnen (Warren & Rosebery, 1990) is another recent effort to identify and to capitalize on how bilingual students learn both content (in this case, science) and language.

Instruction in Mathematics

When considering the quality of instruction, most bilingual-education-program evaluations have focused on the role of the native language or on the role of instruction in developing students' English language skills. In their review of research on the teaching of bilingual learners, Fillmore and Valadez (1986) considered whether mathematical knowledge would transfer from a child's native language into English and when mathematics -- the universal language -- could be taught all in English. The Longitudinal Study documented how teachers dominated classroom conversations and how they asked very low-level questions when they tried to bring their students into a conversation (Ramirez, 1986; Ramirez, Yuen, & Ramey, 1991; Ramirez, Yuen, Ramey, & Pasta, 1991). Ramirez et al.'s critiques of that instruction were based on how such settings are less than optimal for the development of English as a second language. They said nothing about how such settings are also deadly for the development of mathematical knowledge.

The Significant Bilingual Instructional Features Study (Tikunoff, 1985, no date) is the only bilingual-education-program study that I have found to specifically investigate the quality of instruction that LEP students received not just in terms of English-language development, but also in terms of academic development. Tikunoff and his colleagues used models of direct instruction to assess the quality of instruction in bilingual classrooms where native language instruction was in Spanish, Chinese, or Navajo. Unfortunately, their study design commingled mathematics instruction with instruction for other subjects, and it also used time-on-task as a proxy for achievement. Tikunoff's (no date) description of effective instruction in bilingual classrooms is very consistent with -- though not as highly structured as -- Active Mathematics Teaching (Good, Grouws, & Ebmeier, 1983). Beyond direct instruction, Tikunoff and his colleagues identified three teacher behaviors that mediated the effectiveness of direct instruction for LEP students. Effective teachers in Tikunoff's study used both English (L2) and the NES/LES students' native language (L1) for instruction (p. 12). They focused on developing NES/LES students' language, both L1 and L2 (p. 13). And, they responded to and used cultural information during instruction (p. 14). Lending weight to Tikunoff's findings is the fact that direct instruction also has been identified as a characteristic of effective instruction in Chapter 1 settings (Kennedy, Birman, & Demaline, 1986).

Summary Comments

Bilingual-education-program research and evaluation have been driven by concerns for the development of English and of academics among LEP students. These studies have taken for granted the school mathematics curriculum that LEP students are exposed to and, even when problems in instruction are noted, those concerns get cast in terms of language development.

On the one hand, by accepting curriculum and instruction as programmatic given, it has been possible to design and implement evaluations and research studies of increasing sophistication. We really have learned a few things about mathematics teaching for LEP students over the past years, and it would be foolish to pretend that we haven't. It might seem tempting to conclude that we really should continue with business as usual. What is needed, one might be tempted to say, are some better studies that seek to merge mathematics with English-language curricula or that try to document how instruction in mathematics might support the development of language skills. To these efforts, one might recommend adding some attention to closing the achievement gaps, to long-term goals such as advanced coursetaking, and to out-of-school and social goals. But by and large, it might be tempting to not change in any fundamental ways current practices in bilingual-education-program evaluation and research. In the following section, I will argue against such a position. That argument is based on the fact that the general school mathematics curriculum and its teaching have been found wanting on a variety of grounds.

The Shifting Target

Let us assume for a moment that we were able to achieve some of the goals outlined earlier in this paper. Assume that we could close the mathematics-achievement gap between LEP students and their English-proficient peers. Assume further that the gap would remain closed and that these students would enroll in mathematics courses in numbers that were comparable to those of their peers. Though this would be quite an accomplishment, should we be happy with it? If we are to believe the plethora of reports that have come out over the past years, the answer is a resounding NO (American Association for the Advancement of Science [AAAS], 1989; American Mathematical Society [AMS], 1990; Mathematical Association of America [MAA], 1989, 1990, 1991; Mathematical Sciences Education Board [MSEB], 1990; National Council of Teachers of Mathematics [NCTM], 1989, 1991; National Research Council [NRC] 1989, 1991; Steen, 1990). In the event of such success, all that would have been accomplished is that LEP students would be performing at levels

that are judged inadequate when compared to international standards (McKnight, Crosswhite, Dossey, Kifer, Swaford, Travers, & Cooney, 1987; Stevenson, Lummis, Lee, & Stigler, 1990; Stigler, Lee, & Stevenson, 1990). In addition, today's students are encountering insufficient amounts and the wrong kinds of mathematics for what they will need to participate meaningfully in the United States' democratic institutions, in a changing worldwide economic order and its social systems, in the workplace, and for purposes of national security (Secada, 1990b, 1991b; Zarinnia & Romberg, 1987).

The Agenda for Action (NCTM, 1979) argued that problem solving and not the development of basic computational skills should be the focus of school mathematics instruction. Since that time, consensus has been building within the mathematics education community -- comprised of researchers, practitioners, supervisors, and other interested publics -- on a new vision for the content and teaching of school mathematics (Romberg, in press; Romberg & Stewart, 1987). That consensus has been articulated in a series of documents that lay out an agenda for reforming school mathematics in the United States. That agenda is focused on the development of new goals for school mathematics and the development of curriculum, teaching, student assessment, and program evaluation that can support the attainment of these new goals (NCTM, 1989, 1991).

New Goals for School Mathematics

According to the National Council of Teachers of Mathematics (1989), the overarching goal for school mathematics should be the development of a mathematically literate society. For individual students this means the development of

> mathematical power...[or] an individual's abilities to explore, conjecture, and reason logically, as well as the ability to use a variety of mathematical methods effectively to solve non-routine problems....Mathematics [is] more than a collection of concepts and skills to be mastered; it includes methods of investigating and reasoning, means for communication, and notions of context. In addition,...mathematical power involves the development of personal self-confidence. (p.5)

Specifically, the NCTM has proposed five goals for school mathematics. Each student should (1) learn to value mathematics; (2) become[5] confident in her or his abilities to do mathematics, (3) become a mathematical problem solver, (4) learn to communicate mathematically, and (5) learn to reason mathematically (NCTM, 1989, pp. 5-6).

Other writers have approached the specification of more general curricular goals from a different perspective than that of NCTM.

Archbald and Newmann (1988) have written about authentic achievement as that which involves the use of disciplined inquiry to produce knowledge (and a product) that has personal, aesthetic, or social value beyond completing the procedures of school. For authentic achievement in mathematics, goals and school tasks would have to be specified so as to have the aforementioned values that would link mathematics to the world outside of school.

Student Thinking in Mathematics

If curricular goals represent the targets for educational practice, then student thinking is the starting point. The most common criticism of current practice in curricular materials, the content of course coverage, instruction, and assessment is the chasm between how people actually think and learn versus how children are expected to learn in school mathematics. For example, Carpenter (1985) has argued that children enter primary school with much more competence in mathematical reasoning than they are credited with. But, the first-grade arithmetic curriculum, with its stress on memorization of basic facts rather than on problem solving, ignores that competence, and thereby, it lays the groundwork for future school failure. In a later paper, Romberg and Carpenter (1986) built a similar case in criticizing direct instruction for ignoring student thinking.

Similar to writers from within the mathematics education reform movement, Resnick (1987a) has argued that one of the primary functions for schools is teaching students to learn to think. But while writers from mathematics education have chosen examples that are clearly connected to the discipline, Resnick diverges somewhat by drawing on how people learn outside of school (Resnick, 1987b). For example, she describes how knowledge is accumulated and distributed within complex organizations, such as on a large boat, and how individuals have but a portion of the knowledge that is required for the organization to function properly (Resnick, 1987b). Examples like these are more closely aligned with Archbald and Newmann's (1988) notions of authentic learning than the more discipline-based examples found in the NCTM (1989) Curriculum and Evaluation Standards. These different nuances in meaning have implications for teaching and assessment; more on those points later.

Regardless of the disciplinary content of student thinking, there seems to be broad consensus about the nature of that thinking and of learning. Thinking, problem solving, and to some extent learning are thought to share similar characteristics of sense making and of relating new information to established knowledge. Where disagreements occur is in interpretation of the specifics. Information processing models of thinking, for example, require detailed specifications of conditions and of productions that occur under those conditions (e.g.,

Seigler, 1991). The anthropological study of how knowledge is produced, on the other hand, focuses on practices within cultural groups that are thought to create that knowledge and on the social processes by which that knowledge gets validated (e.g., Lave, 1988).

According to information processing and cognitive science theories, knowledge develops in one of three ways: through the gradual accretion of new information to what is already known, through the exposition and resolution of areas of conflict, and through the reorganization of existing knowledge structures. Within the more anthropological traditions, knowledge is thought of as an artifact of human activity. It derives its meaning and validation from that activity and how the activity gets situated within the larger social setting. Hence the processes of knowledge acquisition must be linked to the contexts in which people produce that knowledge.

Many researchers in mathematics education have characterized knowledge as consisting of conceptual and procedural parts (Hiebert, 1986). Conceptual knowledge is interconnected and rich in relations; procedural knowledge produces something. This distinction can be thought of as roughly parallel to the distinction between number concepts (e.g., knowing the concept of 5) and the ability to compute (e.g., knowing how to obtain 2+5). According to Hiebert (1986), mathematics teaching should help students develop and link both sorts of knowledge.

Alternatively, writers who are grounded in information processing models of thinking tend to posit the existence of three broad categories of knowledge: conceptual, procedural, and also strategic (Siegler, 1991). Roughly speaking, one can think of an information processing system as composed of its production rules (procedural knowledge), the conditions that must be met for the system to operate (conceptual knowledge), and an overarching operating system that monitors and regulates the entire process from beginning to end (strategic knowledge). Problem solving consists of the orchestration of all three sorts of knowledge to attain a goal.

Thus even within similar cognition-based approaches to the study of student thinking, there are subtle differences. These differences get played out in different approaches to the specification of curricular tasks, to tracking, and to assessment.

Curricular Tasks and Instruction in Mathematics

If the goals specify the end points, and if student thinking provides the beginnings as well as constraints for school mathematics,

then curricular tasks and instruction should provide the means by which to develop student reasoning and thinking to the desired end points. Again, there is a broad consensus that tasks and instruction should be aligned to the new goals and that they should support the development of student thinking.

Mathematics Curriculum and Tasks

The curriculum has been faulted for failing to produce desired outcomes, for being a disconnected hodgepodge of content, and for lending itself so easily to superficial coverage (Freeman & Porter, 1989; Porter, 1989; Porter, Floden, Freeman, Schmidt, & Schwille, 1988). This lack of cohesion and superficiality do not support the development of conceptual knowledge or of links between conceptual and procedural knowledge (Hiebert, 1986; Romberg & Tufte, 1987). Moreover, this content fails to provide students the disciplinary experiences that they need to develop mathematical power (NCTM, 1989) or the authentic tasks that are necessary for authentic learning to take place (Archbald & Newmann, 1988; Resnick, 1987b).

Hence, new tasks should be developed and organized to provide greater coherence and more depth of coverage (Archbald & Newmann, 1988; Romberg & Tufte, 1987). Those tasks should reflect disciplinary forms as well as authentic forms of mathematical knowledge (Archbald & Newmann, 1988). They should provide students with opportunities to solve problems, to reason mathematically by making conjectures that are then socially validated, to communicate with one another using mathematical language, and to make connections among a variety of representations of the same problem situation (NCTM, 1989). Paper-and-pencil computational facility should be deemphasized; i.e., things like arithmetic algorithms and the solution of algebraic equations through the manipulation of written symbols should be relegated to calculating devices such as calculators and computer software. In place of computations, discrete mathematics, geometry, linear programming, measurement, probability, statistics, and other content should be emphasized (NCTM, 1989).

Some mathematicians go even further in their recommendations for reorganizing the school mathematics curriculum. Steen (1991) and his collaborators would organize mathematics around common themes, like the study of patterns, that cut across and unify seemingly disparate mathematical fields like geometry and statistics. Alternatively, Kaput (1991) has argued for totally scrapping the high school mathematics sequence of Algebra, Geometry, Algebra II, Trigonometry. In its place should be a unified-mathematics sequence that includes new content; relegates all symbolic manipulations to calculators, computers, and other technologies; and uses these technologies to develop depth of understandings and relationships among the different fields of mathematics.

In spite of this agreement on broad goals, there is an emerging tension between the disciplinary and psychological goals of developing mathematical power (NCTM, 1989) versus the criterion that authentic tasks should have external personal, aesthetic, or social value (Archbald & Newmann, 1988). Many tasks found in the mathematics reform documents, while having great disciplinary value, seem to have very little value outside of school. Many tasks that seem very authentic cannot be accomplished within the constraints of the school term, but what is more problematic from a disciplinary point of view, they can be done without reliance on deep mathematical principles.[6]

If mathematics is to be undertaken within rich, real-world problem settings, then another area for debate emerges around the settings that will be chosen for study and therefore will be granted legitimacy as worthy of mathematical scrutiny (Frankenstein, 1989, 1990; Secada, 1991b; Stanic, 1991). In part, this debate revolves around questions of whose interests are served by the study of those contexts and how students are socialized through that study, either explicitly or tacitly (Secada, 1991b). For example, adult students in Frankenstein's (1990) intermediate algebra class learn about percentages by studying how decreasing rates for electricity are linked to increased consumption, and that increased consumption most often entails using appliances that only the wealthy can afford (air conditioners, pool filtration systems, and the like). This analysis of consumption is based on social class. It is in sharp contrast to a mathematical analysis wherein decreasing rates for increased consumption are made to seem as the natural and inevitable outcomes of the so-called laws of supply and demand.

The study of mathematics through authentic contexts also socializes students into accepting certain norms of behavior. For example, a very common activity in elementary school is for students to operate a store of some sort. What seldom, if ever, occurs is for students to run a social-service agency that provides services either for free or on a sliding scale. Presumably one could develop and study exactly the same sorts of mathematical knowledge and skills in either context; yet one context gains legitimacy, the other does not.

Thus, while there is broad-based consensus that mathematics tasks need revamping to support the development of student reasoning, there remain questions about (1) how the new tasks will be organized; (2) the tension between disciplinary knowledge and authenticity; and (3) the cultural contexts that get represented in the curriculum and that thereby will receive legitimacy as being worthy of mathematical study.

Mathematics Instruction

Again there is broad consensus that instruction should support the development of student reasoning, communication, and similar processes that are thought to enhance student learning (Hiebert, in press; Idol & Jones, 1991; Jones & Idol, 1990; Lampert, 1988, 1990a, 1990b; NCTM, 1991). There are some debates about whether or not direct instruction -- as it has been classically understood -- can support student reasoning, especially among students in compensatory programs like Chapter 1 (Brophy, 1991; Collins, 1991; Collins, Hawkins, & Carver, 1991; Idol, Jones, & Mayer, 1991). Some writers who have grounded their analyses of student reasoning from an information processing point of view have argued that direct instruction can incorporate the teaching of specific thinking skills (Idol, Jones, & Mayer, 1991) or that it can include cognitive supports that are slowly withdrawn as students take on increasing responsibility for their own learning (Collins, 1991; Collins et al., 1991). Yet these analyses remove or transform many of direct instruction's defining characteristics -- for example, teachers would no longer directly tell students what they were to learn. Thus, it is not clear that direct instruction as it has been classically understood remains a viable instructional strategy.

Others writers are arguing for a radical overhaul in what constitutes good teaching of mathematics (Ball, 1990; Lampert, 1988, 1990a, 1990b; NCTM, 1991). According to them, teaching is a question of orchestrating student engagement in worthwhile mathematical tasks. A teacher does not tell, but rather he or she poses problems and organizes students into groups to work on those problems. The teacher provides social supports for problem solving, challenges students to justify their responses, and helps students to amplify their justifications when those justifications are not fully developed. The teacher establishes norms of behavior wherein students are to be comfortable participating and are to allow and encourage others to contribute, even when those contributions later do not survive public scrutiny by the whole class.

There are approaches that seem to lie between direct instruction and these more radical departures and they may include features from both. For example, Japanese and other Asian teachers are thought to teach mathematics by spending most of their time on lesson development in whole-class lecture settings. They support student reasoning by discussing one or two problems in great depth, trying to solve them in as many ways as possible. Also, they orchestrate classroom discussion around each student's strategies and try to expose misconceptions as opportunities to revisit and reteach important ideas (Stigler & Stevenson, 1991).

Another approach, known as Cognitively Guided Instruction (CGI), (Carpenter & Fennema, in press; Carpenter et al., 1990; Peterson et al., 1991), combines insights from over five decades of research on how children solve addition and subtraction word problems (Brownell, 1928; Carpenter & Moser, 1984) with more recent research on teacher decision making (Clark & Peterson, 1986). CGI is based on four interlocking assumptions: (1) teachers should know how mathematical content is organized in their children's minds; (2) teachers should make mathematical problem solving the focus of their instruction; (3) teachers should find out what their students are thinking about the content in question; and (4) teachers should make instructional decisions (e.g., the sequencing of topics) based on their knowledge of their students' thinking. Unlike other programs and approaches that prescribe teacher behaviors, this approach relies heavily on teachers' basing their instructional decisions on their knowledge of how their students are thinking about the content (tasks) that they are engaged in.

There are many other issues for instruction, among them, classroom organization. Should the whole class participate in an activity, should it be small groups, or should it be individually based? If instruction is organized by groups, should they be by mathematics ability or heterogeneous? Since mathematics is a social activity, social interaction is necessary. Such interactions are possible not only with small groups but also in whole class settings (e.g., Lampert, 1988, 1990a, 1990b).

Assessment and Evaluation in Mathematics

Goals provide the end point; student cognition, the beginnings and the focus of teaching; and tasks and instruction, the means for achieving those goals. Assessment and evaluation provide evidence that the goals are being met. Assessment focuses on the student; evaluation, on the overall mathematics program in which the student is enrolled.

The NCTM (1989) Curriculum and Evaluation Standards outlined eight aspects of assessment and evaluation, each composed of two poles. One pole should receive emphasis, the other should be deemphasized (p. 191). For example, while assessment should focus on what students know and can do, decreased attention should be placed on what students do not know. Assessment should be ongoing and integral to instruction, not solely for the purpose of assigning grades. And in program evaluation, standardized achievement tests should be one of many possible indicators for monitoring success; other indicators should include samples of student work that are collected in a variety of settings and through a variety of methods. Be-

yond agreement on general principles like these, however, there remain points of debate within fields of both assessment and evaluation.

Assessment

Most simply, argues the NCTM (1989), assessment should be aligned to the new curricula that are intended to achieve newly developing mathematical goals. Indeed, one of the most common complaints about current practice is the failure of tests to be properly aligned to curricula that are in place, even today, and the total misalignment between curricula of the future and present-day tests.

From these misalignments have come two major hypotheses. First is the hypothesis that tests are determining what students actually encounter in their classrooms even if there are broader curricular objectives than are measured by the test (Romberg, Zarinnia, & Williams, 1989; Resnick & Resnick, 1991; Silver, in press). If tests actually are such strong determinants of what gets taught to students (for a counter argument, see Porter, 1989), then current testing practice will derail efforts to reform school mathematics. However, if tests really are such strong determinants of curriculum, then an alternative becomes available; by changing the test, we can change what gets taught (Silver, in press). If we change the tests to include tasks and items that approximate emerging goals for school mathematics, curriculum and instruction will follow. California and Connecticut have adopted this strategy; the former includes open-ended items in its assessment and the latter is committed to using only authentic assessment.

Silver (in press) has argued that this hypothesis may be overly optimistic. Teachers might teach based on a variety of things, not just what is tested -- for example, how they were taught, their beliefs about what constitutes "real" mathematical knowledge, or the press to cover the book. Thus, tests and the curriculum that students are exposed to may be determined by similar forces, but testing per se does not determine the curriculum. Efforts to change curriculum by changing the tests will fail, if not backfire, because they would not address the deeper causes of why teachers teach as they do.

A second hypothesis growing out of the misalignment between testing practice and future curricular needs is a weaker version of the first. Tests, at least standardized achievement tests, are but one of many indicators that teachers rely on in their practice. Because results are so seldom returned to teachers quickly enough or in a format that enables them to make instructional decisions based on the results, standardized tests are, ultimately, unimportant vis-à-vis practice. Their importance lies in their symbolic value, as indicating

a job well done or providing the public with evidence that the schools are working. Teachers attend to tests not because it will help their practice but because it must be done to placate outside interests.

Tests may fail to reveal what students are actually doing in their schools and may interfere with their instruction not because they are dictating curriculum but because they are add-on nuisances.

Paper-and-pencil standardized tests have little utility. As part of an overall reform effort, they need to be changed to support (or at least, not to interfere with) curriculum reform. Be it to mandate or to support curriculum reform, there is consensus: assessment in school mathematics needs revamping (also see Resnick & Resnick, 1991).

Regardless of this consensus, there are still many issues about mathematics assessment that must be worked out. These issues include debates about the kinds of tasks that will comprise these new assessments, the conditions under which they will be completed, what work must be exhibited, scoring rubrics, the creation of performance standards, how to communicate the new rules of testing to participants, and how to communicate the results so that they are meaningful and useful (also see Lajoie, 1991).

Assessment tasks. Beyond agreement that new assessment tasks need developing, there are few exemplars of such tasks and fewer still that would meet the range of criteria found in the various reform documents. An item from the Connecticut assessment reminds students how to compute the volumes of a sphere and of a cone. The task provides a context wherein a scoop of Ben and Jerry's ice cream is placed on a wafer cone. The ice cream forms a perfect sphere of a given diameter. The wafer cone forms a perfect cone of given diameter across the base, with equilateral sides, and is of a given height. The problem is to determine whether the cone could hold all of the ice cream were it to melt.

The samples of some students' work coming from this task are impressive. They clearly understood the need to delve into the mathematical properties of the task for purposes of this assessment. However, this task fails to meet criteria for authenticity as outlined by Archbald and Newmann (1988) or Resnick and Resnick (1991). How could anyone produce a perfect sphere and why would anyone allow a scoop of Ben and Jerry's ice cream to melt -- unless it was for school?[7]

Nominally authentic tasks may fail to reveal the types of mathematical reasoning that are called for in the various documents. For example, students enrolled in an alternative high school conducted surveys of their peers on various topics. They designed, distributed,

and collected the surveys. The students then compiled the results and reported them to the entire school by displaying the results of each month's surveys on the school's bulletin board.[8] On the one hand, it would be very easy (and very authentic) to enter the results of each survey into a software program that would compile them and generate appropriate charts and graphs for display. Such an approach also would close off any opportunity for students to develop and display mathematical competence of the sort that is called for in the reform documents. However, the teacher's intent for this task was to develop mathematical reasoning involving parts of whole, percentages, and the graphical display of data. He did not use the school's readily available computer lab in this activity. Instead, students compiled the results by hand. They converted the results for each question into percentages using calculators, and then they displayed those percentages in pie charts using compass and protractor. In other words, this teacher sacrificed some authenticity in order to develop and to assess student mathematical reasoning.

Authentic assessment tasks are open to the same questions about the standards by which their authenticity is judged as are curricular tasks. For example, the California Assessment (Stenmark, 1989) includes a task wherein students are told that a local college accepts one half of that high-school's graduating class each year while another college also accepts one half of that graduating class. An individual student believes that he is certain to be accepted to one of these two local colleges. The problem is to explain what is wrong with this student's reasoning. At first glance, this task has much out-of-school value, until one realizes that over half of all graduating seniors do not go on to college. One must ask if non-college-intending students would have more than minimal interest in such a task. It seems unlikely that this task will reveal what uninterested students really can do.[9]

One possibility for overcoming problems about cultural and other forms of bias is to allow students to choose among many tasks that include a broad range of cultural contexts, and require comparable mathematical thought, and are to be finished within similar time constraints. For example, a student might enter two raffles for the right to purchase tickets to over-subscribed rock concerts. For the first concert, the odds of winning a pair of tickets would be 50-50; and for the second, the odds could be 60-40. Would this student be assured of getting in to see one of the two shows? Including additional settings increases the likelihood that students will be sufficiently intrigued by at least one of them to actually apply themselves to the task. Some students may actually see the structural parallels among such tasks.

In the past, the search for unbiased test items has meant a search for items that could cut across social class, gender, and cul-

tural categories. One reason that we have such impoverished curricula and tests may be the difficulty in creating such "neutral" tasks. A better strategy might be to create many tasks representing a range of cultural contexts and to ask students to pick the ones that intrigue them the most.

Conditions for assessment. How much time should students have to produce their work? State assessments typically last one to two hours, so that tasks for these assessments would have to be finished within some rather tight time limits.

Such time limits, however, would fail to demonstrate what students could do when engaged in long-term projects. For example, one prototype task developed at the Center on Organization and Restructuring of Schools (CORS)[10] for tenth-grade students provides a setting wherein a family of four moves from Madison, Wisconsin, to the city where the students who are engaged in this task live. The students are given the Thursday and Sunday newspapers of both of these cities since those issues contain information about homes and apartments (for rental or purchase), food, clothing, different kinds of sales, entertainment, job opportunities, and the like. The students are told that, in order to live in Madison, this family spends a certain amount per month that is allocated among these and other budget categories in a certain way. The first problem for the students to solve is: In order to maintain a comparable standard of living, how much per month will this family of four need to spend? Secondly, the students are told to assume that a different group's estimate is double theirs; How would they convince that group that theirs was the right estimate? Third, the students are told that, in order to have so much disposable income, people must earn more since they must pay taxes, social security, health and medical benefits, etc. If in Madison, this family of four's take-home pay was based on a given earned income, what would the earned income have to be in their new home town? And finally, assuming that two people worked in this family, What sorts of jobs would they have to have in order to make ends meet in their new home town? A task like this simply cannot be done in two hours.

Should assessment tasks be uniformly created and administered by an outside agency? Should they be samples of student work that are collected over the course of the year and represent a common core of important tasks as identified by the teacher? Or, should students select their best work and place it into a portfolio that then gets graded? Under current notions of accountability, the first option would be desirable. When issues of curricular validity and alignment are foremost (or when teachers will be evaluated based on their students' work), the second option would seem preferable. However, if one is strictly following models of authenticity -- i.e., what real people do in the real world -- then the last option would be preferred.

Actors, architects, artists, musicians, and even professors up for tenure and promotion assemble their best work for review.

What work should actually be displayed? Testing programs like the California Assessment, the Connecticut Assessment, and many curriculum development projects ask students to show their work enroute to achieving their solutions. This is because these assessments are looking for evidence of mathematical reasoning, communication, and the creation of new knowledge. However, according to standards of authenticity, what should be required are samples of finished work, not the work that was produced while the finished product was being developed. An architect does not include sketches and initial renderings in the final product; musicians do not include rehearsal tapes in their portfolios; business people do not give all the details of why they recommend something in their memos; nor do mathematicians include the false starts in their final articles.

The difficulty with asking for a final product, however, is that it hides the disciplinary work that went into its production. Consider, for example, the task described earlier wherein high school students surveyed their peers and presented the results of those surveys to the school. It is difficult, if not impossible, to determine what mathematical understandings these students actually used in creating those pie charts. For example, we might assume that in the production of these pie charts, students had to have converted individual responses into percentages; i.e., for question number 4, the number of students responding a, b, c, or d would have to be converted into the percent of the students who chose each of these options. We might assume that, in making this conversion, a student had demonstrated knowledge of how parts of a whole are related to percentages.

But consider the case of the student that I observed working on this step of the task. Someone else had already converted the raw scores into percents for all 20 questions on the survey. But she had noticed that the percents for each question did not always add to up to 100. When she pointed this out to her teacher, he told her that not every student had answered every question; for example, 30 of the 31 surveys that had been returned included a response to question 2. Thus, this student was busy checking all of the questions; for those that did not add up to 100 percent, she would divide all of the responses by 30, i.e., by how many people had answered question 2 and not by how many people had actually answered each specific question.

After recomputing the percentages for the questions that needed to be recomputed, she checked her totals and became very distressed when many of them still did not add up to 100! I asked her if she knew why they <u>should</u> add to 100. She responded because she had learned it in another class. Then, I pointed out that, in some cases,

her sums came out to 99 percent and she didn't seem to mind that, but that when the sums came out to 98 percent she got upset. She did not answer that there might have been a rounding error or even that 99 was closer to 100 than 98. Instead, she commented that 99 was just a better answer to get. So next I asked her why she was dividing everything by 30. Her answer was that she had done so for question 2 and in that case the percentages added up to 100.[11]

Since she did not link these questions about parts and whole to how she might resolve her dilemma, I tried to explain to her that a different number of people had responded to every question on the survey. For instance, 30 people had responded to question 2, but only 29 had responded to question 4 and, of those 29, 11 had chosen a. She still did not make the connection that she needed to divide the 11 by 29 because 29 was the appropriate whole for that particular question on the survey. What percent of 29 is 11? I continued. Still no response or indication of understanding. Instead, she kept insisting that the answers had to add up to 100 percent. She did not see how, by dividing the responses for question 4 by 29, she would satisfy this condition. Finally, I suggested that she simply try doing so. Afterwards, I suggested that she add up the percents to this question one more time. Of course, they totaled 100 percent; but, rather than try to understand why this particular example worked, she adduced a general rule -- divide the response to each question by the number of people who actually responded to that question. And very happily, she proceeded to complete the task.

This episode raises many issues in terms of how this student was linking (or failing to link) conceptual understandings about percentages and parts of a whole to the procedures by which she was converting individual student responses to aggregate percentages. However, in this and every other student's final product, there will be no evidence about whether or not such understandings were created, strengthened, or even used. All that remains are the end products of that effort. It is not surprising, therefore, that students are told to show their work in an effort to determine whether they are displaying the forms of mathematical reasoning that the tasks are meant to support.

On the other hand, I have seen samples of work where students were scored lower because the work that they displayed lacked coherence, which is exactly how work in progress is characterized. In one extreme case, I saw a short, concise explanation wherein a student had gone straight to the heart of the task and had done so elegantly. But this work was in a lower corner of the page, lost in the jumble of his other work. We are still struggling to find some middle ground on this issue

Scoring rubrics. As noted earlier in this section, one of the reasons students are asked to show their work is in order for someone else to score the quality of that work against certain standards. The actual content of those standards is still under discussion. Task performance could be scored according to a learning theory or some other criterion (Lajoie, 1991).

In CGI, first grade teachers are taught to assess their students' knowledge on an ongoing basis (Carpenter & Fennema, in press; Carpenter et al., 1990). In those assessments, teachers rely on a well structured body of research on learning wherein a student's right or wrong answers can be linked to how difficult that problem was either in terms of its semantic structures or in terms of the size of the numbers that were used. The strategies that children use when solving various word problems also provide teachers with information about how their students are thinking of the problems. While not written down as formal scoring rubrics, these assessment techniques rely on judgments that are linked to a very rich and detailed specification of how children learn, i.e., to a highly localized learning theory.

Where such specificity is not possible -- in most of the rest of school mathematics -- we could still generate scoring rubrics based on more general learning theories. For example, cognitive scientists (Siegler, 1991) often posit the existence of three kinds of knowledge: conceptual, procedural, and strategic. Lane (1991) included these categories of knowledge in scoring rubrics that were developed for assessing middle school students' performance on a range of authentic tasks.

Alternatively, non-psychological criteria could be developed. Stenmark (1989) describes a general scoring rubric that was used in scoring the open-ended questions of the California Assessment. This rubric was used to score student performance on open-ended questions based on the clarity and coherence of the response; the appropriate use of pictures or diagrams; the quality of the presentation to the intended audience; the use of mathematical reasoning, ideas, and processes; and the nature and flow of the argument that was developed in the response.

If one follows the tenets of CGI and if assessment is supposed to serve instructional purposes, then scoring rubrics should combine explicit learning theories for the tasks at hand with some way of targeting that learning to a coherent end point. While pedagogically these would be the most useful rubrics to develop, they also are the most difficult. We simply do not have as detailed models for how students learn mathematics in domains outside of arithmetic word problems. More general rubrics, like that developed by Lane (1991), may be the best that we can do. The utility of such rubrics for instructional or accountability purposes would remain an open question.

Alternatively, one could create scoring rubrics based on out-of-school models of adequate performance. In some settings, conciseness is more important than the flow of an argument. For portfolios, no learning-theory-based scoring rubric may be adequate, since portfolios are supposed to contain only finished work. Moreover, in the real world, portfolios are scored on a case-by-case basis; i.e., an individual's (or a group's) work is evaluated anew every time that person seeks employment.

Performance standards. To be used, scoring rubrics must contain not just the content or dimensions of interest but also standards against which to judge how people actually perform. How those standards should be developed and calibrated remains an open issue. For example, it is possible to create a priority standards by reference to some absolute criterion or by looking at how experts do the tasks. However, such standards may be set so high that no one had a chance of scoring at the top levels; they might be calibrated in such a way that pedagogically important distinctions got lost, or the experts (if their performance is used) might approach the task in ways that no one else would.

As an alternative, some people recommend that we gather samples of people's work and calibrate the rubrics against those standards. The objection to this, however, is that the performance criteria will end up, essentially, being set too low.

As new cohorts of students become more acclimated to new curricula, new instruction, and these new ways of assessing performance, it is likely that performance will improve. Hence, the performance standards that are settled on will need to be recalibrated every few years. In some sports (ice skating and diving), for instance, performance criteria are recalibrated after someone obtains a perfect score during a major competition.

Performance standards will also need to be linked to instructional practice and to accountability systems. If the standards are calibrated so high or so coarsely that all students cluster around a single level, then they will not be very helpful. On the other hand, if the standards are too finely calibrated, the scorers, teachers, and other consumers of the results may spend so much time trying to understand the distinctions between levels that they will have too little time to use the information for its intended purposes.

The new rules of the assessment game. Under the old rules of testing, students knew pretty much what was expected of them. They either got the answer right or wrong. In the case of teacher-made tests, students know to show enough work to get some partial credit in case the answer is wrong, but not to show so much work that, if the answer is right, they lose credit for work that is wrong or sloppy.

Given the many different ways of scoring performance according to the purposes and the theories that underlie each rubric, it is not clear how students will know exactly what is expected of them. Are they to produce a final, polished product? Should they omit a large amount of detail? Should they include their scratch work? If so, how can they distinguish that work, with all of its false starts, zig-zags, and lack of coherence, from the work that they wish to present? More generally, how does one communicate to a student that he or she is being scored on the use of conceptual knowledge, procedural knowledge, communication skills, or any of the other criteria that have been created? I have not seen anyone grapple with these questions, but they would seem increasingly important, especially for students from diverse backgrounds.

Communication of results. Assessment results are to serve a wide range of purposes. They must be communicated to teachers in ways that will help them make instructional decisions. Ideally such information would combine a description of student competence with some ways of placing that performance along some developmental path. Students, parents, and other stakeholders will also be interested in assessment results. How to report these in ways that all of these interested publics will understand and be able to use remains an open issue.

Program Evaluation

With so much emphasis on assessment, relatively little effort seems to have been placed on program evaluation. In part, because of the belief that, if we first change the assessment, the evaluation systems must change -- if no other way, at least in the sorts of information on which judgments are made.

The NCTM Curriculum and Evaluation Standards (1989) do provide some general suggestions. Evaluation should draw on a wide range of sources of information. Evaluation should focus on ensuring that all students (not just a few) are learning and developing their mathematical power. Evaluations should go beyond looking at student outcome data; they should also focus on the quality of the curriculum in terms of its coherence and content coverage, on the adequacy of materials and other resources, and on the quality of instruction that students receive.

In view of the originally stated purposes for the mathematics reform movement -- participation in the nation's various institutions by the next generation of students -- one long-term outcome should also be evaluated, i.e., do students who experience school mathematics as is recommended by the current reform movement actually participate in our society in the ways that they are expected to?

Summary Comments

The goals for teaching school mathematics have shifted radically. The agenda now is to shift practice to meet those goals. These new goals are focused on the development of students' mathematical power through attention to coherence, depth of study, communication, conjecturing, and the actual doing of mathematics in a variety of contexts that will support such disciplinary inquiry. As the MSEB (1991) so clearly summarized:

> Goals for student performance are shifting from a narrow focus on routine skills to development of broad-based mathematical power (p.5).

Though there are many issues that are still being worked on, there is broad consensus among mathematicians, curriculum developers, psychologists, practitioners, and many key public stakeholders that this shift is necessary because current practice is inadequate.

In current bilingual-education-program-evaluation practice, there are many points that are at odds with how school mathematics is shifting. If we continue to do more of the same, even if we do a better job, we may achieve our goals, but they are outdated and inadequate for purposes of preparing LEP students to participate in the world in which they will live their adult lives. Student assessment, program evaluation, and related research need to shift in order to match these evolving goals. What is more, bilingual education research needs to inform the mathematics education reform movement of what has been learned about the educational needs of LEP students. It is to these points that this manuscript now turns.

Mathematics Education and Bilingual Education: A Two-Way Conversion

On three points I am in total agreement with the current reform movement. First, student thinking and reasoning are the keys to this effort. We are in the business of teaching so that students can develop those skills. However, we still need to unpack what these notions mean vis-à-vis the bilingual learner.

Second, curriculum, instruction, assessment, and evaluation should be coherent, linked, and in support of the development of student thinking. Curriculum and instruction should focus on covering fewer things but providing for greater depth of coverage. Assessment and evaluation should be aligned with and support efforts to teach. They come after everything else, not by themselves.

And third, the goals for education should be linked to the larger society in which our students will live. In some places, the reform movement does not go far enough. We need to consider the situations in which LEP students (and indeed, increasing numbers of our students) find themselves. We should not shy away from the fact that many students live in desperate poverty and that education needs to help them deal with the realities of that as well.

It would be easy to make these general observations and to argue that some immediate and obvious implications for bilingual education grow from them. But the situation is more complex than would be implied by such a one-way conversation. Though changes in school mathematics may have implications for the education of LEP students, bilingual education has much to say to the reformers, not only about the education of LEP student but also about issues that include equity, culture, and performance assessment. Anyone who is even vaguely familiar with the history of bilingual education in this country should have a sense of deja vu when reading about some of the debates in mathematics education. There is much that bilingual education research can say to inform those debates.

Goals

An immediate, albeit not so obvious, implication of the changing goals in school mathematics is that the academic goals for bilingual education need to be reexamined. Academic achievement and advanced course taking should be revisited from the perspective of the kinds of courses that LEP students get placed into. One of the most often told stories for any reform is that people who are positioned to take advantage of it receive a disproportionate amount of the benefits from that change (Secada, 1991b, in press). It is important to monitor how LEP students are included in (or excluded from) reform in schools and districts and at the state and national levels. Beyond vague claims about excellence for all, we need to ensure that inclusion is meaningful.

Elsewhere, the author and others who are concerned about equity in education (Secada & Meyer, 1991) have argued that the mathematics reform movement has not paid adequate attention to these issues. For example, in laying out the reasons for needed reform in school mathematics, the Curriculum and Evaluation Standards (NCTM, 1989) gave the mathematics achievement of minorities as one of the reasons for needing reforms, yet nowhere else within the document does one find specific attention to ensuring that the proposed changes will, in fact, be helpful to minorities. To be fair, in the Teaching Standards (NCTM, 1991) there is a bit (but not that much) more attention paid to equity.

The point may seem like a subtle one. Someone could argue that student diversity need not receive specific and ongoing attention. Absent evidence that LEP students will be omitted or ill served by the reform, such efforts are covered under the rubric of reform for everyone. My counter argument is that silence on issues of student diversity leaves open the very real possibility that, within the reform of school mathematics, stratification of students along the lines of race, social class, language proficiency, or some other means will be recreated. For example, one of the people whose practice is held up as an exemplar for mathematics teaching is Magdalene Lampert (1988, 1990a, 1990b). Anyone who reads her thoughtful analyses of teaching and the skillful ways by which she focuses on student understanding should be impressed by the vision of teaching and the possibilities that she describes. In a recent paper, Lampert (1990a) wrote about her efforts to construct meanings for fractions and computations in her fifth grade classroom. In one particular vignette, she discussed how a community of discourse was formed and maintained in the class.

> Students asserted their contributions and other students revised them. The end result was produced with little teacher input, except asking for clarification and recording on chalk board what was said. All but four members of the class made an active contribution to this discussion; two of the students who did not contribute had very limited English-speaking ability. (p. 263)

In other words, half of the students who were omitted from the community of discourse for this episode were limited English proficient.[12] Note how it seems as if these students' limited English proficiency is the reason for their nonparticipation. Thereby, their exclusion from a discourse community (which is by definition a social fabrication) is made to seem natural and is legitimated.

And this is precisely my point. By their failure to specifically include equity and student diversity as concerns that are integrated from the very start, the various reform documents make possible the restratification of opportunity along the lines by which it has taken place in the past. As the goals get articulated, we must continually ask, Who are the goals for? People in bilingual education must advocate meaningful inclusion of LEP students.

The long-term and out-of-school goals for bilingual education need revisiting. One of the reform movement's main pillars is that today's students need preparation for tomorrow's world -- including access to jobs and meaningful participation in our society. Such goals are commonly missing from similar discussions in bilingual education on grounds that the development of English is the more pressing concern. It would be a major contradiction, however, to argue that the goals for the mathematics education of LEP students

should be linked to out-of-school outcomes but that the goals for the larger program should not.

As part of the concern for social goals, we need to move somewhat beyond the development of mathematical power. We must ask about the knowledge and skills LEP students must have in order to participate meaningfully in American society. There is ample research -- much of it being carried out by people involved in bilingual education -- to document how people are discriminated against due to skin color, accent, and the like. Recently, for example, the Secretary of Labor issued a report that documented the glass ceilings that women and minorities encounter in large U.S. corporations. The question cannot be avoided: What must LEP students know and be able to do in order to overcome those barriers? The answer is likely to include much of what the reform documents say, but it is also likely to diverge in some significant ways.

This general issue is also one that the mathematics reform movement needs to address. Everybody Counts as the NRC (1989) avers. But the question remains, Counts for what purposes? Answers from the bilingual education community should inform a similar debate in the mathematics education community.

The Bilingual Learner of Mathematics

We need to create a view of the LEP student as a learner of mathematics that combines what we know about how mathematics is learned with what we know about second-language learning. Among current efforts that could be helpful in this regard are research on learning strategies (O'Malley & Chamot, 1990), content-based ESL (Crandall, 1987), and the relationship between bilingualism and enhanced functioning in the academic areas (Hakuta, 1986; Secada, 1991a).

It may be helpful to look for common learning processes that cut across language learning and mathematics learning. Such domains might include psychological processes that are common to understanding language and mathematics (Kintsch & Greeno, 1985) as well as for producing either linguistic or mathematical output once someone understands something; sociolinguistic and cultural processes that support the creation of discourse communities in school and how sensemaking takes place and gets validated within such communities (Heath, 1986; Lampert, 1988, 1990a, 1990b; Lave, 1988; NCTM, 1991; Simich-Dudgeon, McCreedy, & Schleppegrell, 1988/89); and how variation in sociocultural contexts affects performance (Stanic, 1991; Zentella, 1981). Of course, distinctions based on content will need to be made; obviously, the retelling or translating of an

arithmetic word problem calls on, at some point, different processes than the solution of that problem.

We need to be careful that the analyses of how bilingual people learn mathematics are not always seen as derivative of research employing monolingual populations. Many analyses are based on the notion that bilingual people are the minority and that research concerning them can be thought of as an application of what we have learned about the majority. This assumption, however, is simply wrong. The norm, within the world, is to be bilingual (Skutnabb-Kangas, 1988).

The research issue is not just the adapting of research concerning monolingual populations to bilingual populations. The more basic research issue concerns the generalizability of results that were found in monolingual populations to the case for bilingual ones. It may well be that much research does generalize. But we cannot tell since we have not developed a unified view of the bilingual learner of mathematics. In a real sense, we are only beginning to learn how sense making occurs in such populations and hence what it means to say that student reasoning -- for the bilingual student -- is the starting point for school mathematics. There is much work to be done.

Curriculum and Instruction

The simplest and most straightforward implication of the mathematics reform movement to the case for bilingual education is that curriculum and teaching for bilingual learners should support the development of their mathematical reasoning. But since we are not clear on the full scope of such a claim, much work still remains in the area of curriculum and instruction.

One promising line of work might be to expand notions that have been found in content-based-ESL and language-learning approaches to create a more unified view of the tasks and instructional methods. It would be helpful to understand where structural analyses of what has become known as the mathematics register (Crandall et al., in press; Dale & Cuevas, 1987; Spanos et al., 1988) diverge from sociolinguistic analyses of communication in classrooms, specifically in mathematics classrooms (e.g., Cazden, 1986; Lampert, 1988, 1990a, 1990b). In the structural analyses, meaning seems somehow to reside in the language and symbols of mathematics. Not surprisingly, direct instruction is used to develop such meanings (e.g., Chamot & O'Malley, 1988). Alternatively, sociolinguistic analyses are more dynamic. They place the development of meaning for symbols within contexts where those symbols are needed to communicate mathematics in meaningful and unambiguous ways -- much as whole

language approaches to reading place the development of vocabulary in context.

There may be more value, from the standpoint of curriculum and teaching, in relying on sociolinguistic as opposed to structural analyses of the mathematics register. Such an analysis would seem more consistent with how Cazden (1986) describes a register as a sociolinguistic construct. Structural analyses of the mathematics register may also increase the fragmentation in the mathematics curriculum for LEP students. Not only are there lessons for skills development but also for mathematics vocabulary and symbolism. This does not mean that structural analyses of mathematical language are not helpful. Indeed, the addition and subtraction problem solving literature relies very heavily on them (Carpenter & Moser, 1984; Secada, 1991a). But a more unified approach would seem, at present, to be called for. It may turn out that attention to higher level structural units -- such as paragraphs, texts, and discourse frames -- will provide greater payoffs than in the past.

The debates on social and cultural referents in mathematics tasks could be informed by similar debates within bilingual education. If mathematics educators are going to take seriously questions of out-of-school outcomes and task authenticity, then they also will need to attend to the situations in which bilingual learners live. Work by Moll, Velez-Ibanez, and Greenberg (1990) in literacy development might provide some ways of proceeding here. A range of social and cultural contexts will need to be represented in newly developing mathematical tasks. We need to develop guidelines for including contexts that are unfamiliar to mainstream cultures and ways for teachers to capitalize on the mathematics that can be learned in such settings.

Newly developing models for teaching mathematics should be scrutinized for their applicability to bilingual learners and adapted as necessary. Lampert's (1990a) acknowledgement of the limitations in her teaching is a reason to question but it is not a reason to reject the developing visions for teaching mathematics (NCTM, 1991). Maybe, with some adjustments -- specifically inviting these students to add their thoughts, encouraging them to use their native languages and asking others to translate, slowing down the fast-paced tempo of the classroom, creating an atmosphere in which language variation in the community of discourse is an accepted fact of life -- these methods can apply to bilingual learners. After all, we should not need to reinvent the wheel for every population.

But also, bilingual educators should develop models for teaching mathematics to bilingual students that are not derivative. Lisa Delpit (1986) wrote about the dilemmas of a progressive black educator having to zig-zag between what seems to be today's faddish way

of teaching and established ways that work for African American students. She wrote about the search for an authentic way of teaching these children that combines what is successful with them with these emerging developments. As these models are developed, they should inform what occurs in school mathematics. Interestingly, the teachers in Cheche Konnen (Warren & Rosebery, 1990) were not certified in science; they were bilingual teachers who must have used their own knowledge of their students to help guide and develop their program. Now that program is being exported, from bilingual classrooms to the entire school.

Assessment and Evaluation

The issues raised earlier vis-à-vis authentic assessment become increasingly complex when they relate to the bilingual learner. There are, of course, some simple techniques in bilingual education for enhancing student understanding of a task. These include rewriting and simplifying language, using familiar contexts, and providing concrete referents. Difficulties will become immediately obvious with the development and application of scoring rubrics and of performance standards.

On one hand, rubrics that are based on learning theories will have to be modified to ensure that evidence concerning actual knowledge of mathematics is obtained and that evidence is not confounded with difficulties that some children may have expressing themselves in English. On the other hand, if unified theories for learning mathematics and a second language could be developed, it might be possible to create tasks and rubrics based on those theories.

Bilingual educators have had much experience in using scoring rubrics that rely on judgments about the quality of linguistic performance, viz., the assessment of oral language proficiency. The Language Assessment Scales (De Avila & Duncan, 1981; Duncan and De Avila, 1986, 1987) include the collection of speech samples, as does the Functional Language Assessment for older students (Hamayan, Kwiat, & Perlman, 1985). Scoring of these samples is against English-speaking norms, which would be the equivalent of calibrating performance on mathematics assessment against expert performance.

It may be possible to create unified assessments that serve multiple purposes. For example, someone might read some mathematics problems to an LEP student and ask the student to repeat each problem before solving it. Student repetitions could serve as speech samples that would be scored along lines of proficiency. Theories of short-term memory for bilingual populations might provide a means for scoring the same sample along lines of what the student under-

stood about the problem. Then, the problem's solution could be scored as an indicator of the student's actual mathematical knowledge. An additional value to such an approach -- besides its cost effectiveness -- is that language proficiency would be assessed using language similar to what the student would encounter in the classroom.

There is reason for concern about the new rules for assessment and culturally diverse populations. There is an increasing body of evidence that children are socialized according to diverse norms when it comes to how performance on socially desirable tasks is evaluated (Deyhle, 1987; Fillmore, 1989, 1990). Deyhle (1987) documented how American Indian children are socialized to judge for themselves when a task has been learned well enough to be put on display and that judgments about performance quality are highly inappropriate. Hence, assessment tasks that ask such students to show all of their work or timed tasks may be met with resistance by some minority-language students.

The NCTM (1989) recommendations for program evaluation are well taken. Outcome data are not adequate for evaluating the quality of the mathematics programs that students encounter. This recommendation takes on particular importance in view of the traditional reluctance for bilingual-education-program evaluation and research to look at the quality of the school mathematics that LEP students encounter. Mathematics educators will need to understand, however, that bilingual-education-program evaluation needs to consider not just the academic aspects of a program but also language development.

Such evaluation efforts would be helped were there to be some clearly articulated theories that look for points where programs can develop both mathematics and English language proficiency (and also the native language, as appropriate), places where one aspect should take precedence, and places where there must be trade offs.

Concluding Comments

Program evaluation is, in part, an issue of asking about effectiveness. One could liken it to asking about a car's gas mileage to see whether it is worth buying. If so, then the evaluation of mathematics programs for bilingual learners in a time of reform is akin to asking not only about gas mileage but also asking for the answer while the car is running and simultaneously being rebuilt from the ground up -- not an easy task.

There is much worth in the current school mathematics reform movement. That assumption is tacit insofar as I refer to the moving

target and am arguing that bilingual education programs need to begin to shift their own goals in light of the new goals for mathematics. Also, there is much of worth in previous bilingual-education-program evaluation. I argue that the conversation needs to go both ways; that people in the education of LEP students should adapt but also should be unafraid of developing ways for teaching the bilingual learner that are not derivative; and that in the history of bilingual education research there have been debates that are similar to those currently found in mathematics education.

We should not think that all debates have been resolved or that most of the technical questions have been answered. Indeed, those efforts are merely beginning. And insofar as there remain open issues and questions, there is room for those who are involved in the education of LEP students to affect that movement through our own practice and research.

Notes

[1] Or, if one follows Baker and de Kanter's (1983) criteria, Does the program work better than any other alternative program?

[2] This report has been very criticized for its many technical flaws (Secada, 1990a; Willig, 1985).

[3] These models are structured-English-immersion strategy, early-exit and late-exit transitional bilingual education programs. They are defined and operationalized in Ramirez (1986) and in Ramirez, Yuen, Ramey and Pasta (1991).

[4] English monolingual students, English-only Hispanics, Spanish-only Hispanics, English-Spanish bilingual Hispanics, Italian-English bilinguals, French-English bilinguals, and German-English bilinguals.

[5] Given Carpenter's (1985) arguments about the knowledge that children enter school with and the results of the program known as Cognitively Guided Instruction (Carpenter, Fennema, Peterson, Chiang, & Loef, 1990; Peterson, Fennema, & Carpenter, in press), maybe this goal should be changed to each student should REMAIN confident in her or his abilities to do mathematics.

[6] In her comments on an earlier draft of this paper, Mary Lindquist raises an additional point. Even the most authentic tasks may suffer from a problem with "so what." For all of our efforts to design such tasks, students (or adults for that matter) may still reject them as uninteresting or as irrelevant. In her comments, for example, Lindquist pointed out that she moved from Madison to someplace else, but she did not engage in sorts of mathematical work that I have proposed as an authentic task elsewhere in this manuscript.

[7] Again, I would like to acknowledge Mary Lindquist's comments on this point. As she notes, part of the power of mathematics comes from our assuming that things are -- for all practical purposes -- like these idealized shapes. We solve the problem in the ideal setting and then apply it to the real world. While granting the need to assume an ideal world -- but only sometimes -- my other objections stand. Who would let a Ben and Jerry's ice cream cone melt all the way? And wouldn't the cone leak anyway?

[8] This graphical representation of real data also has been reported by Warren and Rosebery in Cheche Konnen (1990a, 1990b).

[9] This is not to argue that tasks like these should not serve instructional purposes. Indeed, problems like this one might make college a more viable after high school option for students who seldom, if ever, think of it as an option. While a worthy instructional task, this task is too biased as a stand-alone assessment task to be useful.

[10] I would like to acknowledge Sherian Foster and Matthew Weinstein's contributions to those efforts.

[11] Recall that her teacher had told her to divide by 30.

[12] I would not be so distressed were half of Lampert's class LEP. Then, one could argue that the techniques for creating discourse communities are being invented and refined, and that they do not result in a disproportionate exclusion of students. Lampert does not write about this.

References

American Association for the Advancement of Science (1989). Science for all Americans. Washington, DC: Author.

American Mathematical Society (1990). Mathematicians and education reform (CBMS issues in mathematics education, volume 1). Providence, RI: Author.

Andersson, T. & Boyer, M. (1978). Bilingual schooling in the United States (2nd ed.). Austin, TX: National Educational Laboratory Publishers, Inc.

Archbald, D. A. & Newmann, F. M. (1988). Beyond standardized testing: Assessing authentic academic achievement in the secondary school. Reston, VA: National Association of Secondary School Principals.

Baker, K. A., & de Kanter, A. A. (1983). Federal policy and the effectiveness of bilingual education. In K. A. Baker & A. A. de Kanter (Eds.), Bilingual education: A reappraisal of federal policy (pp. 33-86). Lexington, MA: Lexington Books.

Ball, D. L. (1990). Halves, pieces, and twoths: Constructing and using representational contexts in teaching fractions. East Lansing, MI: National Center for Research on Teacher Education, School of Education, Michigan State University.

Boruch, R. F. & Cordray, D. S. (1980). An appraisal of educational program evaluations: Federal, states, and local agencies (Report to U.S. Department of Education, contract no. 300-79-0467). Evanston, IL: Northwestern University.

Bratt-Paulston, C. (1980). Bilingual education: Theories and issues. Rowley, MA: Newbury House.

Brophy, J. E. (1991). Effective schooling for disadvantaged students. In M. S. Knapp & P. M. Shields (Eds.), Better schooling for children of poverty: Alternatives to conventional wisdom (pp. 211-234). Berkeley, CA: McCutchan.

Brownell, W. A. (1928). The development of children's number ideas in the primary grades (Supplementary educational monograph no. 35). Chicago: University of Chicago Press.

Campbell, D. T. & Stanley, J. C. (1966). Experimental and quasi-experimental designs for research. Chicago: Rand McNally.

Cantoni-Harvey, G. (1987). Content-area language instruction: Approaches and strategies. Reading, MA: Addison-Wesley.

Carpenter, T. P. (1985). Learning to add and subtract: An exercise in problem solving. In E. A. Silver (Ed.), Teaching and learning mathematical problem solving: Multiple research perspectives (pp. 17-40). Hillsdale, NJ: Lawrence Erlbaum.

Carpenter, T. P., Fennema, E., Peterson, P. L., Chiang, C. & Loef, M. (1990). Using knowledge of children's mathematical thinking in classroom teaching: An experimental study. American Educational Research Journal, 26(4), 499-531.

Carpenter, T. P. & Fennema, E. (in press). Cognitively guided instruction: building on the knowledge of students and teachers. International Journal of Educational Research (Special issue on The reform of school mathematics in the United States, edited by W. G. Secada).

Carpenter, T. P. & Moser, J. M. (1984). The acquisition of addition and subtraction concepts in grades one through three. Journal for Research in Mathematics Education, 15(3), 179-202.

Cazden, C. B. (1986). Classroom discourse. In M. C. Wittrock (Ed.), Handbook of research on teaching (2nd ed., pp. 432-463). New York: Macmillan.

Chipman, S. f. & Thomas, V. G. (1987). The participation of women and minorities in mathematical, scientific, and technical fields. In E. Z. Rothkopf (Ed.), Review of Research in Education (Vol. 14, Chapter 9, pp. 387-430). Washington, DC: American Educational Research Association.

Chamot, A. U. & O'Malley, M. (1987). Mathematics book A: Learning strategies for problem solving (Language development through content series). Reading, MA: Addison Wesley.

Clark, C. M. & Peterson, P. L. (1986). Teachers' though processes. In M. C. Wittrock (Ed.), Handbook of research on teaching (2nd ed., pp. 255-296). New York: Macmillan.

Collins, A. (1991). Cognitive apprenticeship and instructional technology. In L. Idol & B. F. Jones (Eds.), Educational values and cognitive instruction: Implications for reform (pp. 121-138). Hillsdale, NJ: Lawrence Erlbaum.

Collins, A. Hawkins, J., & Carver, S. M. (1991). A cognitive apprenticeship for disadvantaged students. In B. Means, C. Chelemer, & M. S. Knapp (Eds.), Teaching advanced skills to at-risk students: Views from research and practice (pp. 216-243). San Francisco, CA: Jossey Bass.

Cook, T. D., & Campbell, D. T. (1979). Quasi-experimentation: Design and analysis issues for field settings. Chicago: Rand McNally.

Cook, T. D., & Shadish, W. R., Jr. (1986). Program evaluation: The worldly science. Annual Review of Psychology, 37, 193-232.

Crandall, J. A. (Ed.) (1987). ESL through content-area instruction: Mathematics, science, social studies. Englewood Cliffs, NJ: Prentice Hall.

Crandall, J. A., Dale, T. C., Rhodes, N. C., & Spanos, G. (1987). English language skills for basic algebra. Englewood Cliffs, NJ: Prentice Hall.

39

Crandall, J. A., Dale, T. C., Rhodes, N. C., & Spanos G. (in press). The language of mathematics: The English barrier. Proceedings of the 1985 Delaware Symposium on Language Studies (VII). Newark: University of Delaware Press. [Available from the authors at the Center for Applied Linguistics, Washington, DC.]

Cronbach, L. J. (and Associates) (1980). Towards reform of program evaluation. San Francisco: Jossey Bass.

Dale, T. C. & Cuevas, G. J. (1987). Integrating language and mathematics learning. In J. A. Crandall (Ed.), ESL through content-area instruction: Mathematics, science, social studies (pp. 9-54). Englewood Cliffs, NJ: Prentice Hall.

Danoff, M. N. (1978). Evaluation of the impact of ESEA Title VII Spanish/English bilingual education program: Overview of study and findings. Palo Alto, CA: American Institutes for Research. [Also available through ERIC Document Reproduction Service.]

Danoff, M. N., Coles, G. J., McLaughlin, D. H., & Reynolds, D. J. (1977-1978). Evaluation of the impact of ESEA Title VII Spanish/English bilingual education program (3 vols.). Palo Alto, CA: American Institutes for Research. [Also available through ERIC Document Reproduction System.]

De Avila, E. A., Cohen, E. G., & Intili, J. K. (1982). Improving cognition: A multicultural approach. In S. S. Seidner (Ed.), Issues of language assessment: Foundations and research. Springfield: Illinois State Board of Education.

De Avila, E. A. & Duncan, S. E. (1981). A convergent approach to oral language assessment: Theoretical and technical specification on the Language Assessment Scales (LAS) Form A (Stock 621). San Rafael, CA: Linguametrics.

De Avila, E. A., Duncan, S. E., Navarrete, C. (1987). Finding out/ Descubrimiento. Northvale, NJ: Santillana.

Delpit, L. D., (1986). Skills and other dilemmas of a progressive Black educator. Harvard Education Review, 56(4), 379-385.

Deyhle, D. (1987). Learning failure: Tests as gatekeepers and the culturally different child. In H. Trueba (Ed.), Success or failure? Learning and the language minority child (pp. 85-108). Rowley, MA: Newbury House.

Duncan, S. E. & De Avila, E. A. (1986). Pre-LAS user's manual (Form A). San Rafael, CA: Linguametrics.

Duncan, S. E. & De Avila, E. A. (1987). Pre-LAS Espanol user's manual (Form A). San Rafael, CA: Linguametrics.

Epstein, N. (1977). Language, ethnicity and the schools: Policy alternatives for bilingual-bicultural education. Washington, DC: George Washington University, Institute for Educational Leadership.

Fillmore, L. W. (1989). Language and cultural issues in early education. Manuscript available from the author, University of California, Berkeley.

Fillmore, L. W. (1990). Latino families and the schools. Remarks prepared for the Seminar on California's Changing Face of Race Relations: New Ethics in the 1990's, Sacramento, CA. Available from the author, University of California, Berkeley.

Fillmore, L. W. & Valadez, C. M. (1986). Teaching bilingual learners. In M. C. Wittrock (Ed.), Handbook of research on teaching (2nd ed., pp. 648-685). New York: Macmillan.

Frankenstein, M. (1989). Relearning mathematics: A different third R--radical maths. London: Free Association Books.

Frankenstein, M. (1990). Incorporating race, class and gender issues into a critical math literacy curriculum. Journal of Negro Education, 59(3), 336-347.

Freeman, D. J. & Porter, A. C. (1989). Do textbooks dictate the content of mathematics instruction in elementary schools? American Educational Research Journal, 26, 403-421.

Good, T., Grouws, D., & Ebmeier, H. (1983). Active mathematics teaching. New York: Longman.

Government Accounting Office (1987). Bilingual education: A new look at the research evidence (Briefing report to the Chairman, Committee on Education and Labor, House of Representatives. GAO/PEMD-87-12BR). Washington, DC: Author.

Hakuta, K. (1986). Mirror of language. New York: Basic Books.

Hamayan, E. V., Kwiat, J. A., & Perlman, R. (1985). The identification and assessment of language minority students: A handbook for educators. Arlington Heights: Illinois Resource Center.

41

Heath, S. B. (1986). Sociocultural contexts of language development. In Bilingual Education Office, California State Department of Education (Eds.), Beyond language: Social and cultural factors in schooling language minority students (pp. 143-186). Los Angeles: Evaluation, Dissemination and Assessment Center, California State University-Los Angeles.

Hiebert, J. (Ed.) (1986). Procedural and conceptual knowledge: The case of mathematics. Hillsdale, NJ: Lawrence Erlbaum.

Hiebert, J. (in press). Reflection and communication: Cognitive considerations in school mathematics reform. International Journal of Educational Research (Special issue on The reform of school mathematics in the United States, edited by W. G. Secada).

Idol, L. & Jones, B. F. (Eds.) (1991). Educational values and cognitive instruction: Implications for reform. Hillsdale, NJ: Lawrence Erlbaum.

Idol, L., Jones, B. F., & Mayer, R. E. (1991). Classroom instruction: The teaching of thinking. In L. Idol & B. F. Jones (Eds.), Educational values and cognitive instruction: Implications for reform (pp. 65-120). Hillsdale, NJ: Lawrence Erlbaum.

Jones, B. F. & Idol, L. (Eds.) (1990). Dimensions of thinking and cognitive instruction. Hillsdale, NJ: Lawrence Erlbaum.

Kaput, J. J. (1991, May 15). Forming a research and development agenda for algebra and quantitative analysis (Discussion draft). Madison, WI: University of Wisconsin-Madison, Wisconsin Center for Educational Research, National Center for Research in Mathematical Sciences Education.

Kennedy, M. M., Birman, B. F., & Demaline, R. E. (1986). The effectiveness of Chapter 1 services. Washington, DC: Government Printing Office.

Kennedy, M. M., Jung, R. K., & Orland, M. E. (1986). Poverty, achievement, and the distribution of compensatory education services. Washington, DC: Government Printing Office.

Kintsch, W., & Greeno, J. G. (1985). Understanding and solving arithmetic word problems. Psychology Review, 92, 109-129.

Lajoie, S. (1991). A framework for authentic assessment in mathematics. NCRMSE Research Review: The Teaching and Learning of Mathematics, 1(1), 6-12. [Available from the National Center for Research in Mathematical Sciences Education, Wisconsin Center for Education Research, School of Education, University of Wisconsin-Madison]

Lampert, M. (1988). The teacher's role in reinventing meaning of mathematical knowledge in the classroom (Research series no. 186). East Lansing, MI: Institute for Research on Teaching, College of Education, Michigan State University.

Lampert, M. (199a). Connecting inventions with conventions. In L. P. Steffe & T. Wood (Eds.), Transforming children's mathematics education (pp. 253-265). Hillsdale, NJ: Lawrence Erlbaum.

Lampert, M. (1990b). When the problem is not the questions and the solution is not the answer: Mathematical knowing and teaching. American Educational Research Journal, 27(1), 29-64.

Lane, S. (1991). The conceptual framework for the development of a mathematics assessment instrument for QUASAR. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago.

Lave, J. (1988). Cognition in practice. New York: Cambridge University Press.

Lindblom, C. E. & Cohen, D. K. (1979). Usable knowledge: Social science and social problem solving. New Haven, CT: Yale University Press.

Mackey, W., & Beebe, V. N. (1977). Bilingual schools for a bicultural community. Rowley, MA: Newbury House.

Mathematical Association of America (1989). Reshaping college mathematics. Washington, DC: Author.

Mathematical Association of America (1990). Challenges for college mathematic: An agenda for the next decade. Washington, DC: Author.

Mathematical Association of America (1991). A call for change: Recommendations for the mathematical preparation of teachers of mathematics. Washington, DC: Author.

Mathematical Sciences Education Board, National Research Council (1990). Reshaping school mathematics: A philosophy and framework for curriculum. Washington, DC: National Academy Press.

Mathematical Sciences Education Board, National Research Council (1991). Counting on you: Actions supporting mathematics teaching standards. Washington, DC: National Academy Press.

McKnight, C. C., Crosswhite, F. J., Dossey, J. A., Kifer, E., Swafford, J. O., Travers, K. J., & Cooney, T. J. (1987). The underachieving curriculum: Assessing U.S. school mathematics from an international perspective. Champaign, IL: Stipes.

Mohan, B. A. (1986). Language and content. Reading, MA: Addison Wesley.

Moll, L. C., Velez-Ibanez, C., & Greenberg, J. (1990). Community knowledge and classroom practice: Combining resources for literacy instruction (A technical report from the Innovative Approaches Research Project). Tucson, AZ: College of Education, University of Arizona. [Also available from Development Associates, Arlington, VA]

Myers, D. E. & Milne, A. M. (1982, October). Mathematics opportunity and exposure: Language minority status and gender (Draft. Prepared for Part C: Math and Language Minority Students Project, U.S. ED contract no. 300-80-0778). Washington, DC: Decision Resources.

National Council of Teachers of Mathematics (1979). Agenda for action. Reston, VA: Author.

National Council of Teachers of Mathematics (1989). Curriculum and evaluation standards for school mathematics. Reston, VA: Author.

National Council of Teachers of Mathematics (1991). Professional standards for teaching mathematics. Reston, VA: Author.

National Research Council (1989). Everybody counts: A report to the nation on the future of mathematics education. Washington, DC: National Academy Press.

National Research Council (1991). Moving beyond myths: Revitalizing undergraduate mathematics. Washington, DC: National Academy Press.

Oakes, J. (1990a). Opportunities, achievement, and choice: Women and minority students in science and mathematics. In C. B. Cazden (Ed.), Review of Research in Education, 16, 153-222.

Oakes, J. (1990b). Multiplying inequalities: The effects of race, social class, and tracking on opportunities to learn mathematics and science. Santa Monica, CA: RAND.

O'Malley, J. M., & Chamot, A. U. (1990). Learning strategies in second language acquisition. New York: Cambridge University Press.

Orum, L. S. (1983). The question of effectiveness: A blueprint for examining the effects of the federal bilingual education program. NABE News, 6, 3-13.

Orr, E. W. (1987). Twice as less. New York: Basic.

Peterson, P. L., Carpenter, T. P., & Fennema, E. (1991). Using children's mathematical knowledge. In B. Means, C. Chelemer, & M. S. Knapp (Eds.), Teaching advanced skills to at-risk students: Views from research and practice (pp. 68-101). San Francisco: Jossey Bass.

Porter, A. (1989). A curriculum out of balance: the case of elementary school mathematics. Educational Researcher, 18(5), 9-15.

Porter, A., Floden, R., Freeman, D., Schmidt, W., & Schwille, J. (1988). Content determinants in elementary school mathematics. In D. A. Grouws, T. J. Cooney, & D. Jones (Eds.), Effective mathematics teaching (pp. 96-113). Hillsdale, NJ: Lawrence Erlbaum.

Ramirez, J. D. (1986). Comparing structured English immersion and bilingual education: First-year results of a national study. American Journal of Education, 95, 122-149.

Ramirez, J. D., Pasta, D. J., Yuen, S. D., Billings, D. K., & Ramey, D. R. (1991). Final report: Longitudinal study of structured English immersion strategy, early-exit and later-exit transitional bilingual education for language-minority children. Volume II (Contract no. 300-87-0156). San Mateo, CA: Aguirre International.

Ramirez, J. D., Yuen, S. D., & Ramey, D. R. (1991). Final report: Longitudinal study of structured English immersion strategy, early-exit and later-exit transitional bilingual education for language-minority children. Executive summary (Contract no. 300-87-0156). San Mateo, CA: Aguirre International.

Ramirez, J. D., Yuen, S. D., Ramey, D. R., & Pasta, D. J. (1991). Final report: Longitudinal study of structured English immersion strategy, early-exit and later-exit transitional bilingual education for language-minority children. Volume I (Contract no. 300-87-0156). San Mateo, CA: Aguirre International.

Reauthorization of expiring federal elementary and secondary education programs: Bilingual education (Vol. 4. Hearing before the subcommittee on elementary, secondary, and vocational education of the Committee on Education and Labor, House of Representatives. Hearing held in Washington, DC, March 24, 1987. Serial no. 100-5). Washington, DC: Government Printing Office.

Resnick, L. B. (1987a). Education and learning to think. Washington, DC: National Academy Press.

Resnick, L. B. (1987b). Learning in school and out. Educational Researcher, 16(9), 13-20.

Resnick, L. B., & Resnick, D. P. (1991). Assessing the thinking curriculum: New tools for educational reform. In B. R. Gifford & M. C. O'Connor (Eds.), Changing assessments: Alternative views of aptitude, achievement, and instruction (pp. 37-75). Boston: Kluwer.

Romberg, T. A. & Carpenter, T. P. (1986). Research on teaching and learning mathematics: Two disciplines of scientific inquiry. In M. C. Wittrock (Ed.), Handbook for research on teaching (2nd ed., pp. 850-873). New York: Macmillan.

Romberg, T. A., & Stewart, D. (Eds.) (1987). The monitoring of school mathematics (3 vols.). Madison: University of Wisconsin-Madison, Wisconsin Center for Education Research, School of Education.

Romberg, T. A., & Tufte, R. (1987). Mathematics curriculum engineering: Some suggestions from cognitive science. In T. A. Romberg & D. Stewart D. (Eds.), The monitoring of school mathematics (Vol. no. 2, Chapter 15, pp. 71-108). Madison: University of Wisconsin-Madison, Wisconsin Center for Education Research, School of Education.

Romberg, T. A., Zarinnia, A., & Williams, S. (1989). The influence of mandated testing on mathematics instruction: Grade 8 teachers' perception. Madison: University of Wisconsin-Madison, Wisconsin Center for Education Research, National Center for Research in Mathematical Sciences Education.

Secada, W. G. (1990a). Research, politics, and bilingual education. Annals of the American Association on Political and Social Science, 508, 81-106 (Special issue on Bilingual Education edited by C. Cazden and C. Snow).

Secada, W. G. (1990b). The challenges of a changing world for mathematics education. In T. J. Cooney (Ed.), Teaching and learning mathematics for the 1990s. (1990 Yearbook of the National Council of Teachers of Mathematics pp. 135-143). Reston, VA: NCTM.

Secada, W. G. (1991a). Degree of bilingualism and arithmetic problem solving in Hispanic first graders. Elementary School Journal, 92(2), 211-229.

Secada, W. G. (1991b). Agenda setting, enlightened self-interest, and equity in mathematics education. Peabody Journal of Education, 66(2), 22-56. [Special issue on Needed: An agenda for equity in mathematics education, edited by W. G. Secada & M. R. Meyer]

Secada, W. G. (in press). Race, ethnicity, social class, language, and achievement in mathematics. In D. Grouws (Ed.), Handbook of research on mathematics teaching and learning (pp. 623-660). New York: Macmillan.

Secada, W. G. & Meyer, M. R. (Eds.) (1991). Needed: An agenda for equity in mathematics education [Special issue]. Peabody Journal of Education, 66.

Siegler, R. (1991). Children's learning (2nd ed.). Englewood Cliffs, NJ: Prentice Hall.

Silver, E. A. (in press). Assessment and mathematics education reform in the United States. International Journal of Educational Research (Special issue on The reform of school mathematics in the United States, edited by W. G. Secada).

Simich-Dudgeon, C., McCreedy, L., Schleppegrell, M. (1988/89). Helping limited English proficient children communicate in the classroom. A handbook for teachers (Program Information Guide Series no. 9). National Clearinghouse for Bilingual Education.

Skutnabb-Kangas, T. (1988). Multilingualism and the education of minority children. In T. Skutnabb-Kangas & J. Cummins (Eds.), Minority education: From shame to struggle (pp. 1-44). Clevedon, England: Multilingual Matters.

Spanos, G., Rhodes, N. C., Dale, T. C., & Crandall, J. (1988). Linguistic features of mathematical problem solving: Insights and applications. In R. R. Cocking & J. P. Mestre (Eds.), Linguistic and cultural influences on learning mathematics (pp. 221-240). Hillsdale, NJ: Lawrence Erlbaum.

Stallings, J. A. & Stipek, D. (1986). Research on early childhood and elementary school teaching programs. In. M. C. Wittrock (Ed.), Handbook of research on teaching (2nd ed., pp. 727-753). New York: Macmillan.

Stanic, G. M. A., (1991). Social inequality, cultural discontinuity, and equity in mathematics. Peabody Journal of Education, 66(2), 57-71. [Special issue on Needed: An agenda for equity in mathematics education, edited by W. G. Secada & M. R. Meyer]

Steen, L. A. (Ed.) (1990). On the shoulders of giants: New approaches to numeracy. Washington, DC: National Academy Press.

Steen, L. A. (1991). Mathematics for all Americans. In T. J. Cooney (Ed.), Teaching and learning mathematics in the 1990s (1990 Yearbook of the National Council of Teachers of Mathematics, pp. 130-134). Reston, VA: NCTM.

Stein, C. B., Jr. (1986). Sink or swim: The politics of bilingual education. New York: Praeger.

Stenmark, J. K. (1989). Assessment alternatives in mathematics: An overview of assessment techniques that promote learning. Berkeley, CA: University of California, EQUALS, Lawrence Hall of Science.

Stevenson, H. W., Lummis, M., Lee, S-Y., & Stigler, J. W. (1990). Making the grade in mathematics: Elementary school mathematics in the United States, Taiwan, and Japan. Reston, VA: National Council of Teachers of Mathematics.

Stigler, J. W., Lees, S-Y., & Stevenson, H. W. (1990). Mathematical knowledge: Mathematical knowledge of Japanese, Chinese, and American elementary school children. Reston, VA: National Council of Teachers of Mathematics.

Stigler, J. W., & Stevenson, H. W. (1991, Spring). How Asian teachers polish each lesson to perfection. American Educator, 12-20, 43-47.

Tallmadge, G. K., Lam, T. C. M., Gamel, N. N. (1987a). The evaluation of bilingual education programs for language-minority, limited-English-proficient students: A status report with recommendations for future development (Phase I report). Mountain View, CA: RMC Research Corporation.

Tallmadge, G. K., Lam, T. C. M., Gamel, N. N. (1987b). Bilingual education evaluation system: Users' guide. (2 vols.). Mountain View, CA: RMC Research Corporation.

Tikunoff, W. J. (1985). Applying significant bilingual instructional features in the classroom. Washington, DC: National Clearinghouse for Bilingual Education.

Warren, B. & Rosebery, A. S. (1990). Cheche Konnen: Collaborative scientific inquiry in language minority classrooms (2 vols.). Cambridge, MA: Technical Education Research Association. [Also available from Development Associates, Arlington, VA]

White, S. H., & Buka, S. L. (1987). Early education: Programs, traditions, and policies. Review of Research in Education, 14, 43-91.

Willig, A. C. (1985). A meta-analysis of selected studies on the effectiveness of bilingual education. Review of Educational Research, 55, 269-317.

Zarinnia, E. A., & Romberg, T. A. (1987). A new world view and its impact on school mathematics. In T. A. Romberg & D. A. Stewart (Eds.), The monitoring of school mathematics (Vol. 1, pp. 21-61). Madison: Wisconsin Center for Education Research, University of Wisconsin.

Zentella, A. C. (1981). Ta bien: You could answer me in cualquier idioma. Puerto Rican codeswitching in bilingual classrooms. In R. Duran (Ed.), Latino language and communication behavior (pp. 109-131). Norwood, NJ: Ablex.

Response to Walter Secada's Presentation

Penelope L. Peterson
University of Michigan

Following up on Walter's comments, I just have to say that all you need to know about me can really be summarized by the fact that I was on the faculty at the University of Wisconsin--Madison for 11 years, and I left. I am now at Michigan State University, quite happily by the way, although my years at Madison were very productive.

At Michigan State University, we have really an outstanding and very interesting group of scholars working on problems of reform of teaching and teacher education in schools. Much of what I am going to say and my ideas and my thinking have been influenced profoundly by my conversations with my colleagues within this community of learners, teachers, and researchers that we have created in the College of Education at Michigan State University. Specifically, I would like to acknowledge the contributions to my own thinking and learning that have resulted from ongoing conversations over the last five years with Deborah Ball, David Cohen, Patrick Dickson, Magdalene Lampert, Sarah McCarthey, Richard Prawat, Ralph Putnam, and Suzanne Wilson.

Our Dean, Judy Lanier, has been influential in creating this thoughtful community of learners, teachers, and scholars in our college. And so I would like to start out with a little metaphor that Judy Lanier has used to talk about this whole problem of assessment and to raise questions about the idea that many people have that "assessment will drive instruction." Judy questions this drive to construct a national test to measure the progress toward reform in education in our nation's schools. Lanier compares our race toward reform with our race to make it to the moon in the 1960s, and she queries: By designing a national test to measure the progress of reform, isn't it a bit like setting the goal to make it to the moon, designing a terrifically big new telescope to see if we made it there, but doing nothing in between?

A major message of Walter Secada's paper is that there is a lot "in between" that needs to be considered seriously. There is a lot in between that we need to think about and take account of if we are really to measure and understand education change and progress. We need to think hard about some of the elements that Secada has pointed out.

I would like to situate my remarks within the context of reform in mathematics education because mathematics education is really

the context for Walter's remarks and for Mary Lindquist's comments as well. And, as Walter points out in his paper, we in the mathematics education community are perceived to have a coherent vision for reform. This vision encompasses and extends from the standards of the National Council of Teachers of Mathematics (NCTM) (1989; 1991) to include ongoing reform efforts of members and affiliates of that organization. But it also encompasses the mathematics education reform efforts that are going on in states such as California with the California Mathematics Framework (1987; 1992) that proposes a new and ambitious vision of mathematics instruction. In their remarks, both Walter Secada and NCTM President, Mary Lindquist, have done a nice job of summarizing that vision.

That vision interweaves four important elements. One is that there are new goals for students' learning of mathematics that move beyond computation. The second element is a significant revision in the K-12 mathematics curriculum -- new topics are added, and others are eliminated or reduced. Third, this reform vision really call for a different kind of pedagogy. An important idea is that how mathematics is taught shapes what students learn. Consequently, the reform proposals call for students to talk much more and teachers to talk less, for students to make conjectures and arguments, and for teachers to skillfully direct and moderate students' investigations. Finally, the proposals call for attention to the mathematics learning of all students, African-American, Hispanic, and female students as well as white males.

Now to pick up on that and to quote from Walter's paper: "The shifting of goals and visions for school mathematics has profound implications for the education of LEP students. Assume, for example, that we actually achieved the goals for mathematics that are found at least tacitly in current evaluation practices. Would this be a real success, or would it not be a pyrrhic victory? Were we to succeed in meeting the mathematics goals that are found in current tests, LEP students would become computational wizards, but would be unable to engage in the sorts of mathematical activities that their English-proficient peers would engage in routinely during their own schooling. The target has shifted: the evaluation of school mathematics for LEP students needs to shift as well. Conversely, the mathematics reform movement has failed to pay serious attention to the education of diverse learners....Unfortunately, the new Standards for school mathematics curriculum and its teaching do not include checks to ensure that they will, in fact, apply to everyone, and that resultant practice will meet the diverse needs of this country's LEP students."

I think that one thing that is clear from Walter's paper and from Mary's remarks is that the problems with mathematics instruction are systemic, and that achievement of these ambitious goals will require changes in curriculum, assessment, policies, and structures at

all levels of the system from the state to the district to the school to the classroom. What I would like to focus on here is what I see as the invisible actor in Secada's paper, but perhaps the key person in systemic reform -- the teacher. What I have to say is intended to embellish on the arguments that Walter has made in his paper.

In Secada's concluding comments, he uses an apt metaphor to reveal a major difficulty that reformers face. Secada contends that "the evaluation of mathematics programs for bilingual learners in a time of reform is akin to asking not just about a car's gas mileage [to see whether it is worth buying], but asking new questions and asking for the answer while the car is running and simultaneously being rebuilt from the ground up -- not an easy task."

Not an easy task, I agree, but a very apt metaphor. In fact, a similar metaphor that we are fond of using at Michigan State is the idea that, as educators involved in reform at all levels, we are trying to sail a boat while we are building it -- the same idea as driving the car while you're building it from the ground up. But, as I read this metaphor in Secada's paper (and maybe it's because the focus of my research has always been on teachers and teaching and these are what I spend my life thinking about) I just kept thinking -- but it all depends on who is driving the car. What is missing for me in this metaphor is the driver who is driving the car while rebuilding it. The most important driver right now in our American schools and in our nation's classrooms is the teacher. And, what I would like to spend my fifteen minutes talking about is the teacher because I think without teacher support, without active participation on the part of teachers, without profound changes in teacher's beliefs, knowledge, thinking, understanding and expectations, little is going to change. Teachers are the critical mediators of student's mathematics learning, and teachers are the critical agents of this reform.

But teachers are in a difficult position, a very difficult position indeed; and anything I say today is not meant in any way to berate teachers. On the contrary, what I think we need to do is figure out how to help and support teachers. Teachers face incredible challenges. Take, for example, the case of mathematics. Teachers are products of the kinds of classrooms that are currently under fire. The mathematics education reforms invite teachers to construct quite a different kind of teaching and learning, yet they themselves never experienced that kind of mathematics teaching and learning. Further, teachers have not experienced the kind of mathematics that reformers are talking about them teaching. It is unclear whether any of us have ever experienced that. Remember, again, this is a car we're rebuilding as we're driving it along or a boat that we're constructing as we try to sail it.

It is a profound dilemma for the teacher in the classroom, and I would like to propose that it is a profound dilemma for us. I want to spend some of the rest of my time talking about an actual case of a teacher -- an authentic case. I would like to tell you a story about Cathy Swift, a California teacher whom I have been following for three years. I would argue that Cathy Swift is typical of many teachers out there and, because of that, we need to try to understand what she has been going through.

In addition to my personal judgment that Ms. Swift is typical, I have statistical data that placed Cathy Swift among a modal cluster of teachers when we surveyed 493 elementary teachers in California, Florida, and Michigan about their current goals and activities in teaching mathematics. (See Peterson, Putnam, Vredevoogd, and Reineke, in press). Cluster analysis of teachers' survey responses yielded five clusters of teachers: (a) primary teachers who had students use manipulatives extensively; (b) *Math their Way* teachers who had students use manipulatives and discuss problem solving extensively; (c) modal teachers whose profile reflected a softened version of drill-and-practice teachers; (d) drill-and-practice teachers; and (e) teachers in the expert cluster whose profile represented a balanced version of the *Math their Way* teachers' profile. Cathy Swift's survey response fell into the modal cluster of teachers. After we conducted this survey, we began to do case studies of twenty-four elementary teachers in the state of California in which we went into their classrooms and interviewed the teachers and observed their mathematics teaching practice.

These case studies are part of a longitudinal study of policy and practice that I have been conducting with several Michigan State colleagues in which we have been examining the relationship between the state level reform in mathematics in California and classroom practice. Building on the notion of systemic reform, the California mathematics education reform has several elements. One element is the California Mathematics Framework (California State Department, 1985; 1992) which lays out the new vision of mathematics, learning, and teaching aimed at "teaching mathematics for understanding." The second element is the selection of textbooks or the design of curriculum materials aligned with the Framework. A third element is the construction of new assessments of students' mathematics learning that are aligned with the Framework and the texts. In our study, we are interested in what teachers are doing when one looks behind the classroom door. Our picture of what we found in teachers' classrooms came out in the Fall, 1990, issue of Educational Evaluation and Policy Analysis (EEPA) in which we provided case studies of five different elementary teachers' classrooms in three different California school districts. (See Ball, 1990; Cohen, 1990; Peterson, 1990; Wiemers, 1990; and Wilson, 1990).

The teacher that I wrote about in my EEPA case study is a teacher whom I call Cathy Swift. Cathy Swift is teaching in a school district that has 118,000 students. It is a very large urban district. The large urban elementary school in which Ms. Swift teaches has an extensive minority population; many immigrants come into this school; many of the students are Limited English Proficient; and most of the students (90 percent) qualify for free or reduced lunch. Substantial ethnic and linguistic diversity exists within the school with 20 different languages being spoken by children who are enrolled. Signs posted in the building and information for families in the staff lounge are in English, Spanish, Lao, Vietnamese, Cambodian, and Hmong. In my initial case study, I summarized my impressions of what I saw as "a smoothly and swiftly-paced model lesson in the tradition of effective teaching for basic skills -- warm-up, review, and seatwork, with continuous monitoring by the teacher, direct instruction, and directed prompting when a student needs help." In other words, I saw Cathy enact marvelous direct instruction lessons in the tradition of active mathematics teaching.

That was in the 1988-89 school year. In my case analysis, I argued that one reason that Cathy taught the way she did was because she was teaching within a model that the school district had adopted called the Achievement for Basic Skills (ABS) model. This model was based on master learning ideas where teachers were given pacing charts, mastery tests to assess students, and additional worksheets to use for remediation when students failed to pass the mastery tests. Teachers were told to use direct instruction, and they had to turn in their pacing charts and their scores on their mastery tests to a mentor teacher in their school who reported them to the ABS office in the district. I argued that Ms. Swift's practice was framed by having to teach within that context. Now I would like to tell you what I saw in Cathy Swift's classroom the following year when I went.

When I returned to Ms. Swift's classroom, it was the 1989-90 school year, and Swift had elected to switch to teaching a group of students that were limited English proficient. She had a class called a "sheltered" class. Although none of her students had English as their native language, Cathy was supposed to teach the class in English, and she did. She taught in a small bungalow that had been added to the school because the school was overcrowded, having been built for 300 students and now housing more than 900 students. When I entered the bungalow, I was struck by Ms. Swift's class -- thirty-one faces looked up at me that varied in shades from yellow to brown to black. The three white faces in the room were Cathy and I and the thirty-second student who was a fair-skinned white girl with bright red hair who was a native Russian speaker. Although Cathy herself speaks no languages other than English, she told me that she had decided to teach this fourth grade sheltered class because she wanted to get out of the "restrictiveness" of the ABS model, and

teachers of sheltered classes were not required to follow the ABS model. Immediately, I thought to myself: "Good! Great! We're going to see interesting mathematics teaching now, right? Fantastic kinds of things."

Cathy began by telling me that what she thought LEP students need is lots of "hands-on" experiences with mathematics, and they need a lot of "active involvement." As I watched Cathy Swift teach a lesson to her LEP students, I saw her attempt to put her ideas into practice. She began with a short review that dealt with "fact families," and then, for the second part of the lesson, she read a book to her students, How much is a million? Ms. Swift read the book aloud and asked her students factual questions that dealt with information in the text. But what was striking was the missed opportunity for asking the students some very interesting questions, such as asking the students to speculate about the size of a million or querying them about what they thought a person might buy with a million dollars.

The last part of Ms. Swift's lesson was, in her words, "a review of place value." Now pretend you were in this classroom situation, and you were sitting there trying to make sense of what was going on, and I will describe to you what was happening was the following. Ms. Swift passed out different colored cards to her fourth-graders. Each card had a number from 0 to 9 written on it, and each child got two cards. The color of each card matched one of the colors of the "places" on the board: the ones' place on the board was beige, the tens' place was pink, the hundreds' place was red, and the thousands' place was blue.

Ms. Swift began the activity by announcing: "I'm going to write a number on the board, and you look at your card. If you have the card that goes in that place, I want you to get up and stand in that place." To demonstrate what she meant, Ms. Swift wrote the number "100" on the board. She wrote a "1" above the red hundreds' place on the board, a "0" above the pink tens' place, and a "0" above the beige ones' place on the board. Then she called on the person with "one hundreds" to come up. Hector announced that he had it so he marched to the board and stood under the red hundreds' place holding his card in front of him. Hector was holding up a dark blue card with a one on it.

Ms. Swift said to Hector, "No, you have the thousands, not the hundreds." Holding up Hector's card, she asked the class, "Does this go in the hundreds place?"

The class chorused in unison, "No!"

Ms. Swift said, "Then, well, who has the hundreds' place?"

One child called out, "the red."

The child with the "1" on a red card came up and stood beneath the red one hundreds' place on the board.

Ms. Swift then asked, "Now, what do we have in our tenth place?"

The class chorused in unison, "zero!"

The teacher queried, "Who has that one?" and a child with a pink card with a "0" on it came up and stood beneath the tens' place at the board.

Finally, Ms. Swift asked, "Okay, who has the ones' place?"

A girl with a beige card with a "0" on it went to the board and stood under the ones' place.

Looking at all three children holding their cards at the board, Ms. Swift summarized, "Okay, reds are hundreds, pinks are tens, and beige is the ones' place. Who can read our number for us?" She called on Belinda who responded correctly, "one hundred."

Ms. Swift continued the place value activity for several minutes by having the students enact each of several more numbers. As with the above example, the students were "actively involved" in this "hands-on" activity" as the students with the appropriate cards came to the board to represent the places in the number.

Let me summarize what I see as significant in this case of Cathy Swift -- a teacher who is trying very hard to teach mathematics for understanding to her limited English proficient students. Cathy Swift is a thoughtful, hard-working teacher, and a sensitive, compassionate, caring person. She chose to teach in this large, urban overcrowded school with children from a diversity of ethnic, linguistic, and socioeconomic backgrounds; she could have chosen to teach in a less challenging situation. Looking at Swift's teaching from one perspective of where she was the previous year, she has made significant changes. She has moved beyond the direct instruction model and is engaging in activities that are very much consonant with the mathematics education reform. We saw Cathy attempt to integrate literature into her mathematics teaching by reading a story about numbers to her LEP students. Her students appeared engaged throughout the reading. Further, Cathy is using what she sees as "active involvement" and "hands-on manipulatives" in her mathematics teaching. Cathy thinks of her LEP students as achieving concrete understanding of place value through the kinesthetics of pairing the

placement of their body on the color of their card with where they place the number. Yet from the perspective of most mathematics educators, Cathy's understanding and her practice reflect a rather rote conception of place value.

Writers of the California Mathematics Framework and Model Curriculum Guide (California State Department, 1987) would argue that Ms. Swift has really missed the "essential understanding" of place value which they articulate as follows:

"Any number can be described in terms of how many of each group there are in a series of groups. Each group in the series is a fixed multiple(the base of the place value system) of the next smaller group. The place value system requires the act of counting groups as though they were single items. It is this organizational structure that gives us the power to deal with large numbers and small numbers in reasonable ways. Rather than endless, unfathomable series of numbers, we need only the digits zero to nine. By grouping we can think of a hundred as a unit or a trillion as a unit; by subdividing we can think either of one thousandth or one millionth of a unit. We can record very large and very small numbers by using the position of the digit to indicate the group we are using as a unit" (California State Department of Education, 1987, p. 19).

Why did Cathy Swift teach place value the way she did to her fourth-grade class of LEP students? One way of thinking about Swift's practice is that when she ceased to work within the direct instruction model, she was freed from constraints, but she was also left to recreate her classroom practice from her own knowledge, beliefs, and understandings. So what did Cathy Swift do? She attempted to bootstrap up from her knowledge and understandings which she herself admits are incomplete in the area of mathematics. For example, when I asked Cathy about her mathematics course at the liberal arts college she attended, she said, "it was a joke." Cathy Swift acknowledges that she does not know how to teach children to solve problems. Yet like all of us, what Cathy sees and understands is framed within and limited by her own understandings and perspectives so that she sees only what she can see from her own point of view. So when I asked her about the California Mathematics Framework, Cathy said that she had attended a seminar where they "read the framework from cover to cover." "Great! I thought to myself," so I asked out loud, "What did you think about it? Did you have any new insights?" Swift replied, "Well, actually it's a pretty boring, dull document. I guess it just reaffirms what I'm already doing."

Why do I tell this story of Cathy Swift? I tell it to illustrate for you the average teacher's dilemmas within the contexts of this current education reform. Although I have used this one case, I do believe that, in several important ways, Cathy Swift represents the

typical elementary teacher. In this case, Swift has moved to teaching LEP students so she faces even greater challenges than the typical teacher of white, middle-class students. Cathy Swift's dilemmas are these: she is being asked to teach a new mathematics that is different from the mathematics she learned; with a new pedagogy that is different from the way she was taught; to achieve new goals different from basic skills; to a new group of students, more diverse than those with whom she attended school and who have certainly more diverse ethnic, linguistic, and social knowledge, backgrounds and experiences than Cathy's own reflect. Cathy Swift is being asked to do all this without being supported and helped to attain the kinds of new knowledge and skills that she will need to do it.

I would argue that these are dilemmas that we cannot just let the Cathy Swifts of the world confront alone. We must confront them as well. As teacher educators, policy makers, administrators and researchers, we must somehow confront these dilemmas with Cathy. If we do not confront these dilemmas and help and support teachers in developing the new knowledge, skills, understanding, and dispositions that they will need to reconstruct the car or build the boat, then we will not need to spend millions of dollars to do a meaningful evaluation of mathematics education of limited English proficient students of the kind that Walter Secada so eloquently described in his paper. We can just reread the research reports of the evaluations that have been done over the last decade. We will not need to do a million-dollar evaluation because if we do not join in confronting teachers' dilemmas with them, then nothing significant will change in the mathematics education of the average American student let alone in the education of the average limited English proficient student.

References

Bali, D. L. (1990). Reflections and deflections of policy: The case of Carol Turner. Educational Evaluation and Policy Analysis, 12(3), 247-259.

California State Department of Education (1985). Mathematics framework for California public schools: Kindergarten through grade twelve. Sacramento, CA: Author.

California State Department of Education (1987). Mathematics Model Curriculum Guide. Sacramento, CA: Author.

California State Department of Education (1992). Mathematics framework for California public schools: Kindergarten through grade twelve. Sacramento, CA: Author.

Cohen, D. K. (1991). A revolution in one classroom: The case of Mrs. Oublier. Educational Evaluation and Policy Analysis, 12, 311-30.

National Council of Teachers of Mathematics (1989). Curriculum and evaluation standards for school mathematics. Reston, VA: Author.

National Council of Teachers of Mathematics (1991). Professional standards for teaching mathematics. Reston, VA: Author.

Peterson, P.L. (1990). Cathy Swift: Doing more in the same amount of time. Educational Evaluation and Policy Analysis, 12(3), 261-280.

Peterson, P.L., Putnam, R.T., Vredevoogd, J, & Reineke, J. (in press). Profiles of practice: Elementary school teachers' views of their mathematics teaching. In W. G. Secada (Ed.), Researching educational reform: The case of school mathematics in the United States. (Special issue of the International Journal of Educational Research).

Wiemers, N. J. (1990). Transformation and accommodation: A case study of Joe Scott. Educational Evaluation and Policy Analysis, 12(3), 281-292.

Wilson, S. (1990). A conflict of interests: The case of Mark Black. Educational Evaluation and Policy Analysis, 12(3), 293-310.

Response to Walter Secada's Presentation

Mary Lindquist
Columbus College, Georgia

As the President-elect of the National Council of Teachers of Mathematics (NCTM), let me first thank you for the opportunity to attend this conference and to learn from you. In the few minutes that I have this afternoon, I would like to assist Walter Secada--not that he needs much assistance--in presenting the mathematics community's view toward reform in mathematics, to raise a few concerns related to his paper, and to reinforce Walter's discussion of needed steps in collaboration between you and those of us in mathematics education.

First let me say I could not agree more with one premise of Walter's paper, that "evaluation of school mathematics for LEP students needs to shift," and with one of his warnings, and I quote: "if we continue to do more of the same, even if we try to do a better job of it, we may achieve our goals, but they are outdated and inadequate for purposes of preparing LEP students to participate in the world in which they will live their adult lives." In these quotes, all students could be substituted for LEP students...it is not just your problem...it's a problem for all our students. This is why NCTM responded and produced two documents, the Curriculum and Evaluation Standards for School Mathematics and the Professional Standards for Teaching Mathematics.

Let me share some of my views of the vision of these two documents even though Walter has done an excellent job of explaining the view of the mathematics community. I was afraid when he said it was his critical day -- once in a while he gets real critical on lots of issues -- but he was very gentle today. To consider the mathematics community position, return with me to Thomas Popketwitz's talk this morning when he compared the field of change to a baseball field. I have often felt that we in mathematics education have been an outfield; I hope we have made it to shortstop now. Hopefully, this vision of ours is not just a field of dreams. But if there is a field of dreams, you will come and we together can make a difference.

Walter Secada stated the five goals of the NCTM Standards. Let me reiterate them quickly. Students should become mathematical problem solvers, they need to learn to reason mathematically, and to communicate mathematics. The other two goals address the value of mathematics. Do our students value mathematics? Do the students you work with value mathematics? One result from NAEP: eighth graders across the nation as a whole think mathematics is extremely

important. When asked, important for whom, however, individual students respond: it's important for somebody else, not for me. If we can achieve this last goal, each child should become confident in his or her ability to do mathematics, we will make progress.

Let me say a few things about equity because Walter does address this issue in the paper and make a confession. Early in my teaching career, I thought I had made it when I got five boys, yes, boys, in calculus. For a long time, many of us in mathematics thought of math as a filter, an exclusive club for only a few. One of the changes today is a relook at that attitude. We have made progress, but we have a way to go. Mathematics is still a filter.

As I have looked at National Assessment of Educational Progress (NAEP) data over the years, it has concerned me greatly that we have not provided the opportunity for everybody to experience a broad curriculum. We have closed the gap of performance on numbers and operations among different groups of students. But the gap still exists in measurement, geometry, and problem solving. It's not because our students can't learn; it's because a lot of them are not given the opportunity to learn.

The standards consider this broader view of mathematics. That's one reason we have an algebra standard in K-4. It opens the door to everybody rather than make algebra a cut-off. We need to do more than teach mathematics, year after year, that can be done with a $3.95 calculator. We need more math, and we need different mathematics. I think one of the most exciting aspects of the new vision is the emphasis on communication. At each group of grade levels (K-4, 5-8, and 9-12), there is a standard on communication in mathematics. These are standards that each of you may want to read because they do tie our two interests together.

I want you all to think for a minute of the computation exercise (you might want to write it down) 5 3/4 divided by half. What do our students do with that? There are many of our students that give us an answer. But does it make sense? Can they give you a situation -- I don't know if this is authentic or not -- but can they even give a situation that includes any language other than five and three over four divided by half. What meaning does it have? When they get an answer, does it make sense?

I know the first thing many students say is "five and three quarters, I have to change that to an improper fraction and now what do I do? I think I do something in a circle with those fractions. Divided by half, I think I flip something." There's no meaning there. There's no language there that gives meaning. I believe that almost all of our students across the nation could solve this problem if it were set in a context. Think about it yourself. If you had, and I know that I

won't pick the right context, five and three-fourths pies. If you gave a half a pie to each family, to how many families could you give? Well, think about one pie. If you were going to give half to each family, you'd give to how many families? Two. How about two pies? Four. You all are bright, try five. Ten. Gosh, I'm dividing but I get an answer larger than I started with. That doesn't fit the conception of division held by a lot of our children.

Also, you begin to realize that what you did was multiply by two. Maybe there is something to that flipping or inverting. We need to work with the language, and we need to begin with the children's language and build the mathematic language from their language. In summary, my vision of the standards include mathematics that makes sense to all children.

In Walter's paper he talked about a discipline-based task versus an authentic task. I don't think there is a need to be polar. I think we need both. Mathematics is a discipline; we can't leave the mathematics out of our assessment.

I want to examine two examples that he gives in his paper. One example was about moving from Madison. I moved from Madison once, and I never went through all that. How authentic is that problem for 10th graders. I agree wholeheartedly that we have often taught math so students would do better in the next grade. That's ridiculous. We need to have real life, whatever that is. But remember that real life for young children is often fantasy, and for older children it's not <u>our</u> real life. We do need to make mathematics useful or authentic, but we cannot ignore the discipline. I do not mean to return to the 1930s when all mathematics had to be based on use. If you couldn't use it immediately, it was not included. Mathematics is a discipline, and there's some beautiful mathematics that can excite children and help them look at the discipline itself.

I want to argue a little bit with what Walter says about the Connecticut example. Walter says it's not authentic. Let me give you the task: you have an ice cream cone; on top of that cone you put a scoop of ice cream; if the ice cream is a perfect sphere and it all melted down into the cone, would the cone run over? From the students's responses that I have heard, they don't think that's authentic either, but they play along with us, they get engaged in it, and they come up with a variety of ways to solve the problem.

But Walter made one statement that really, really bothered me. He said it was not authentic, because no scoop of ice cream is ever spherical. But that's what we do in math. We make assumptions; that's the basis for the whole discipline of mathematical modeling. We try to simplify the world so that we can work with it. If I assume it's spherical, that's my mathematical assumption to help me work

the problem. That assumption doesn't bother me at all. That doesn't make it non-authentic.

What makes it non-authentic to me is, "who cares?" What I would rather solve -- I mean, who ever worried about that problem? Loving ice cream, I would rather know how large a scoop I can get on the top without it falling out. You'll understand these references as you read Walter's paper; it's very readable and very good.

I think we have to be careful not to change one set of problems for another. Here I will sound a little bit defensive because Walter says in his paper that the NCTM Standards really have no authentic examples, no uses -- I may be overstating it a little bit -- no part of the world outside. Yet, as I look through it again, I see many recommendations for collecting data, analyzing data, starting with children's own problems, estimating change, making dog kennels, and so forth. There are efforts to tie mathematics to the outside world and to see its usefulness.

Let me quickly change to assessment since that's been the topic of this conference. I think Walter has put assessment in its proper perspective. There is guidance in the Standards: assessment of students, assessment of teaching, and program evaluation. The main focus is on student assessment that is to improve learning and teaching. The emphasis is on what students can do instead of what they cannot do.

As I work with teachers, some of the most exciting things have happened when they interview their students. At first they're amazed the students can't do and don't understand. What bothers me is that I have not always been able to turn that view around so they can tell me what the students can do. When we do get to that stage, they know what to do next.

I think one issue that Walter raises, whether assessment is the driving force, is crucial. Read the quotes from Ed Silver in Walter's paper about the position that changing assessments will not necessarily change learning and teaching. One of the main issues concerns beliefs and expectations. Until teachers change their beliefs, until society changes its beliefs about mathematics, we will not make progress. You know, it's very acceptable in our nation not to be able to do math. Think about it. Until we change that, I don't know that we'll move forward.

In conclusion, I want to comment on the five steps that Walter recommends taking. I think there are steps that you could take alone, but hopefully we will take together in looking at mathematics. I may be paraphrasing some of these, but I think this is what Walter was saying in his paper.

First of all, set goals for LEP students in mathematics that are in concert with NCTM standards. This does not mean they have to be exactly the same but that you are reaching for the new vision of mathematics.

Second, communicate and continue to do research that will inform the mathematics reform about LEP students. We need to know what you are thinking and what your research is saying. I would add that we also need to work together on the research.

Third, develop samples of contexts that may be unfamiliar to the mainstream culture and ways for teachers to use these. You are the ones that can inform curriculum developers and teachers.

Fourth, help wrestle -- and these were not Walter's words -- help wrestle with assessment issues, especially issues regarding language in cultural context. We need that help in mathematics.

Fifth, encourage program evaluations that focus on the quality of school mathematics that students encounter.

As Walter said, the target's moving. But I think if we work together, we have a much better chance of hitting it than if we work separately.